V Jornadas de Ingeniería Telemática

Libro de ponencias

Editor: José J. Pazos Arias

Vigo
12 a 14 de septiembre, 2005

V Jornadas de Ingeniería Telemática

Libro de ponencias

Editor
José J. Pazos Arias

UNIVERSIDAD DE VIGO

Presentación

Alcanzada la V edición de las Jornadas de Ingeniería Telemática (JITEL), nos congratula reconocer que la celebración bianual de este encuentro se ha consolidado como el foro de referencia para el intecambio de las ideas, los progresos técnicos y las experiencias de la comunidad dedicada a la ingeniería telemática en el ámbito nacional, con aportaciones provenientes tanto desde el mundo académico como, cada vez más, del entorno empresarial, algo que debe ser saludado con plena satisfacción. Esta edición, el número de comunicaciones recibidas ha llegado hasta 148, con origen en prácticamente todas las universidades españolas donde existen grupos de investigación en este campo, lo que constituye un claro signo del grado de interés que suscita el evento, así como de la repercusión que va adquiriendo con las sucesivas ediciones. Por otra parte, y por primera vez, se celebran conjuntamente con el JITEL unas Jornadas de Seguimiento de los resultados de 28 proyectos de investigación en el ámbito de la ingeniería telemática financiados por el Plan Nacional de I+D. Con todo ello, la dimensión que ha cobrado el JITEL ha supuesto una labor de organización e intendencia de cierto calibre que, desde el Comité Organizador, confiamos en haber resuelto diligentemente y con discreción. Al mismo tiempo, creemos que supone un fuerte incentivo hacia una mayor apertura de la convocatoria y en favor de una más amplia divulgación de la investigación que en ella se presenta.

Estas Jornadas constituyen, claro está, una empresa colectiva a la que contribuyen, en primer lugar, los autores de las ponencias quienes, con trabajos de una calidad merecedora de encomio, han hecho de la labor del Comité del Programa a la hora de seleccionar las comunicaciones una tarea tan repleta de responsabilidad como grata. El Comité de Programa aprecia también el esfuerzo de los autores cuyas ponencias no han tenido cabida en el programa final pues, aunque privados por distintos motivos de presentar sus resultados, su participación ha contribuido al éxito de la reunión. Hemos de manifestar, asimismo, un caluroso agradecimiento a todos los voluntarios que han prestado servicio como revisores de los artículos porque de su entusiasmo, precisión y puntualidad es consecuencia, casi tanto como del mérito de los autores, el

excelente nivel del programa. Como en otras ediciones anteriores, dentro de las jornadas se celebran dos mesas redondas, cuyos participantes han acogido la invitación con interés y, a buen seguro, su participación resultará enriquecedora para todos. A este respecto, como Presidente del Comité Organizador me gustaría dejar constancia de mi agradecimiento al profesor Julio Berrocal Colmenarejo por su valiosa ayuda en la configuración de la mesa redonda dedicada al análisis de la relación universidad-empresa. Las conferencias invitadas completan el programa de las jornadas, para cuya organización fue inestimable la ayuda de los profesores Arturo Azcorra Saloña y Carlos Delgado Kloos.

Es bien sabido que la organización de un congreso supone la realización de un conjunto pesado de tareas organizativas, poco agradecidas aunque imprescindibles para su funcionamiento. A este respecto, aprovecho para mostrar mi gratitud a los profesores Manuel Ramos Cabrer y Manuel Fernández Veiga quienes han realizado un importante trabajo en la puesta en marcha y atención del servidor web del congreso.

Por último, en nombre del Comité Organizador, reciban también nuestro agradecimiento las instituciones que han patrocinado la celebración de las Jornadas: la Universidad de Vigo, la Xunta de Galicia y el Ministerio de Educación y Ciencia.

Es nuestro deseo que las Jornadas resulten estimulantes y provechosas para todos. Confiamos también en que continúen siendo, en venideras ediciones —la inmediata con los compañeros de Málaga—, tan fecundas y vivificantes como hasta la fecha. Poco arriesgamos al augurarlo, pues las Jornadas son ya una incipiente tradición y un hito para cuantos en ellas hemos participado.

José J. Pazos Arias
Presidente del Comité de Programa
Presidente del Comité Organizador
Vigo, julio de 2005

Comité de programa

Javier Aracil Rico	Universidad Pública de Navarra
Arturo Azcorra Saloña	Universidad Carlos III de Madrid
Julio Berrocal Colmenarejo	Universidad Politécnica de Madrid
Víctor M. Carneiro Díaz	Universidade da Coruña
Vicente Casares Giner	Universitat Politècnica de València
Carlos Delgado Kloos	Universidad Carlos III de Madrid
Yannis A. Dimitriadis	Universidad de Valladolid
Guillem Femenias Nadal	Universitat de les Illes Balears
Antonio Fernández Anta	Universidad Rey Juan Carlos
Julián Fernández Navajas	Universidad de Zaragoza
Sebastián García Galán	Universidad de Jaén
Víctor Guillermo García García	Universidad de Oviedo
Joan García Haro	Universidad Politécnica de Cartagena
Ana Gómez Oliva	Universidad Politécnica de Madrid
Antonio Gómez Skarmeta	Universidad de Murcia
José Luis González Sánchez	Universidad de Extremadura
Klaus Hackbart	Universidad de Cantabria
Xavier Hesselbach Serra	Universitat Politècnica de Catalunya
Eduardo Jacob Taquet	Euskal Herriko Unibertsitatea
Juan Manuel López Soler	Universidad de Granada
Pedro Merino Gómez	Universidad de Málaga
José Juan Pazos Arias	Universidade de Vigo (Presidente)
Álvaro Suárez Sarmiento	Universidad de Las Palmas de Gran Canaria
Juan Ramón Velasco Pérez	Universidad de Alcalá
Joan Vinyes i Sanz	Universitat Pompeu Fabra
Juan Manuel Vozmediano Torres	Universidad de Sevilla

Comité organizador

Luis E. Anido Rifón
Manuel Fernández Veiga
Francisco J. González Castaño
José J. Pazos Arias (Presidente)
Manuel Ramos Cabrer

Revisores

Agüero Calvo, Ramón, U. de Cantabria
Aguilar Igartúa, Mónica, U. Pol. de Catalunya
Alonso Martín, Pedro, U. Pol. de Madrid
Álvarez Díaz, Manuel, U. de La Coruña
Álvarez-Campana Fernández-Corredor, Manuel U. Pol. de Madrid
Amor Pinilla, Mercedes, U. de Málaga
Anido Rifón, Luis E., U. de Vigo
Arco Rodríguez, José M., U. de Alcalá
Arias Fisteus, Jesús, U. Carlos III
Asensio Pérez, Juan I., U. de Valladolid
Bagnulo Braun, Marcelo, U. Carlos III
Banchs Roca, Albert, U. Carlos III
Barba Martí, Antonio, U. Pol. de Catalunya
Bellalta, Boris, U. Pompeu Fabra
Bellas Permuy, Fernando, U. de La Coruña
Bellido Triana, Luis, U. Pol. de Madrid
Bernardos Cano, Carlos J., U. Carlos III
Berrocal Colmenarejo, J., U. Pol. de Madrid
Blázquez del Toro, José M., U. Carlos III
Boronat Seguí, Fernando, U. Pol. de Valencia
Bote Lorenzo, Miguel L., U. de Valladolid
Burguillo Rial, Juan C., U. de Vigo
Cacheda Seijo, Fidel, U. de La Coruña
Calderón Pastor, María, U. Carlos III
Campo Vázquez, Celeste, U. Carlos III
Cano Escribá, Juan C., U. Pol. de Valencia
Carral Pelayo, Juan A., U. de Alcalá
Casares Giner, Vicente, U. Pol. de Valencia
Catalá Monzó, Ángel, U. Pol. de Valencia
Centeno González, José, U. Rey Juan Carlos
Costa Montenegro, Enrique, U. de Vigo
da Silva Fariñas, Antonio, U. Pol. de Madrid
de Miguel Moro, Tomás, U. Pol. de Madrid
Delgado Kloos, Carlos, U. Carlos III
Díaz Redondo, Rebeca P., U. de Vigo
Díaz Verdejo, Jesús E., U. de Granada
Dimitriadis, Yannis, U. de Valladolid
Esparza Martín, Óscar, U. Pol. de Catalunya
Espinosa Acereda, Koldo, U. del País Vasco
Estepa Alonso, Antonio, U. de Sevilla
Estévez Ayres, Iria, U. Carlos III
Fernández Anta, Antonio, U. Rey Juan Carlos
Fernández Cambronero, D., U. Pol. de Madrid
Fernández García, Norberto, U. Carlos III
Fernández Jiménez, Francisco J., U. de Sevilla
Fernández Navajas, Julián, U. de Zaragoza
Fernández Panadero, M^{a}Carmen, U. Carlos III
Fernández Veiga, Manuel, U. de Vigo
Fernández Vilas, Ana, U. de Vigo
Ferrer Gomila, Josep L., U. de les Illes Balears
Ferro Vázquez, Armando, U. del País Vasco
Forné Muñoz, Jordi, U. Pol. de Catalunya
García Arranz, Marta, U. de Cantabria
García Duque, Jorge, U. de Vigo
García Escallé, Pablo U. Pol. de Valencia
García Fernández, Roberto, U. de Oviedo
García Galán, Sebastián, U. de Jaén
García García, Carlos, U. Carlos III
García García, Víctor G., U. de Oviedo
García Gutiérrez, Alberto E., U. de Cantabria
García Hernando, Ana B., U. Pol. de Madrid
García Martínez, Alberto, U. Carlos III
García Moros, José, U. de Zaragoza
García Pañeda, Xabiel, U. de Oviedo
García Reinoso, Jaime, U. Carlos III
García Roger, David U. Pol. de Valencia
García Rubio, Carlos, U. Carlos III
Garijo Ayestarán, Mercedes, U. Pol. de Madrid
Gazo Cervero, Alfonso, U. de Extremadura
Gil Castiñeira, Felipe, U. de Vigo
Gil Solla, Alberto, U. de Vigo
Goirizelaia Ordorika, Iñaki, U. del País Vasco
Gómez Montenegro, C., U. Pol. de Catalunya
Gómez Oliva, Ana, U. Pol. de Madrid
Gómez Sacristán, Ángel, U. Pol. de Valencia
Gómez Sánchez, Eduardo, U. de Valladolid
Gómez Skarmeta, Antonio F., U. de Murcia
González Castaño, Francisco J., U. de Vigo
González Cristóbal, José C., U. Pol. de Madrid
González Martínez, Carlos, U. Pol. de Madrid
González Sánchez, José L., U. de Extremadura
González Vidal, Francisco, U. Pol. de Madrid
Guerrero López, Carmen, U. Carlos III
Guirado Puerta, A. M., U. Pol. de Cartagena
Gutiérrez González, Ll., U. Pol. de Catalunya
Hackbart, Klaus, U. de Cantabria
Hernández Díaz, Vicente, U. Pol. de Madrid
Hernández Solana, Ángela, U. de Zaragoza
Hesselbach Serra, X., U. Pol. de Catalunya
Huecas Fernández-Toribio, Gabriel, U. Pol. de Madrid
Ibáñez Fernández, Guillermo, U. Carlos III
Irastorza Teja, José Á., U. de Cantabria
Izal Azcárate, Míkel, U. Púb. de Navarra
Jacob Taquet, Eduardo, U. del País Vasco
Larrabeiti López, David, U. Carlos III
Lázaro, Jorge, U. Pol. de Valencia
Llamas Nistal, Martín, U. de Vigo
Lloret Mauri, Jaime, U. Pol. de Valencia
López Ardao, José C., U. de Vigo
López Carmona, Miguel Á., U. de Alcalá
López de Vergara, Jorge E., U. Aut. de Madrid
López García, Cándido, U. de Vigo
López Muñoz, Francisco J., U. de Málaga
López Nores, Martín, U. de Vigo
López Santidrián, Lourdes, U. Carlos III
Luque Centeno, Vicente, U. Carlos III
Machado Sánchez, S., U. Pol. de Catalunya
Macías López, Elsa M., U. de Las Palmas de Gran Canaria
Madinabeitia Luque, Germán, U. de Sevilla
Magaña Lizarrondo, E., U. Púb. de Navarra
Malgosa Sanahuja, J., U. Pol. de Cartagena
Manzanares López, P., U. Pol. de Cartagena
Marrero Marrero, Domingo, U. de Las Palmas de Gran Canaria
Marsá Maestre, Iván, U. de Alcalá
Martí, Ramón, U. Pompeu Fabra
Martín Escalona, Israel, U. Pol. de Catalunya
Martín Faus, Isabel V., U. Pol. de Catalunya
Martínez Arrúe, Ignacio U. Pol. de Valencia
Martínez Bauset, Jorge U. Pol. de Valencia
Martínez Cruz, Javier, U. de Málaga
Martínez Martínez, A. E., U. Aut. de Madrid
Martínez Mas, Antonio, U. Pol. de Madrid

Martínez Pérez, Gregorio, U. de Murcia
Martínez Zaldívar, F. J., U. Pol. de Valencia
Marzo Lázaro, José L., U. de Girona
Mata Díaz, Jorge, U. Pol. de Catalunya
Matellán Olivera, Vicente, U. Rey Juan Carlos
Melendi Palacio, Daniel, U. de Oviedo
Merino Gómez, Pedro, U. de Málaga
Miguel Nieto, Carlos, U. Pol. de Madrid
Mompó Gómez, Rafael, U. de Valladolid
Montoto Castelao, Paula, U. de La Coruña
Moreno Blázquez, Jesús, U. Pol. de Madrid
Muñoz Calle, Francisco J., U. de Sevilla
Muñoz Expósito, José E., U. de Jaén
Muñoz Organero, Mario, U. Carlos III
Muñoz Tapia, José L., U. Pol. de Catalunya
Mut Puigserver, Macià, U. de les Illes Balears
Oliver Riera, Miguel, U. Pompeu Fabra
Ortega Daza, Juan J., U. de Málaga
Ortiz, Roberto, U. de Cantabria
Paricio García, Álvaro, U. de Alcalá
Pavón Gómez, Santiago, U. Pol. de Madrid
Pavón Mariño, Pablo, U. Pol. de Cartagena
Payeras Capellà, M., U. de les Illes Balears
Pazos Arias, José J., U. de Vigo
Pegueroles Valles, Josep, U. Pol. de Catalunya
Pérez Belleboni, Emilia, U. Pol. de Madrid
Pinto Alarcón, Mónica, U. de Málaga
Pla Boscá, Vicent U. Pol. de Valencia
Portilla Figueras, José A., U. de Alcalá
Quintana Suárez, Miguel Á., U. de Las Palmas de Gran Canaria
Ramos Cabrer, Manuel, U. de Vigo
Ramos Muñoz, Juan José, U. de Granada
Ramos Pascual, Francisco, U. Pol. de Valencia
Raposo Santiago, Juan, U. de La Coruña
Regueras Santos, Luisa M., U. de Valladolid
Rendón, Juan, U. Pompeu Fabra
Rincón Rivera, David, U. Pol. de Catalunya
Robles Valladares, Tomás, U. Pol. de Madrid
Rodero Merino, Luis, U. Rey Juan Carlos
Rodríguez Cayetano, Manuel, U. de Valladolid
Rodríguez Hernández, Pedro S., U. de Vigo
Rodríguez Pérez, F. J., U. de Extremadura
Rodríguez Rubio, Raúl F., U. de Vigo
Romeral Ortega, Ricardo, U. Carlos III
Rubio Arjona, Lorenzo, U. Pol. de Valencia
Rubio Cifuentes, Gregorio, U. Pol. de Madrid
Ruiz Martínez, Pedro M., U. de Murcia
Ruiz Mas, José, U. de Zaragoza
Ruiz Piñar, Francisco J., U. Pol. de Madrid
Salazar Riaño, José L., U. de Zaragoza
Sallent Ribes, Sebastià, U. Pol. de Catalunya
Salvachúa Rodríguez, J., U. Pol. de Madrid
Sánchez Aarnoutse, J., U. Pol. de Cartagena
Sánchez Fernández, Luis, U. Carlos III
Sánchez Pastor, F. J., U. Aut. de Madrid
Sánchez Rodríguez, David, U. de Las Palmas de Gran Canaria
Sanz Gil, Roberto, U. de Cantabria
Sedano Ruiz, Marifeli, U. Pol. de Madrid
Seepold, Ralf, U. Carlos III
Seoane Pascual, Joaquín, U. Pol. de Madrid
Serrano Yáñez-Mingot, Pablo, U. Carlos III
Soto Campos, Ignacio, U. Carlos III
Suárez González, Andrés, U. de Vigo
Suárez Sarmiento, Álvaro, U. de Las Palmas de Gran Canaria
Traver Salcedo, Vicente, U. Pol. de Valencia
Unzilla Galán, Juan J., U. del País Vasco
Ureña Pascual, Manuel, U. Carlos III
Valdovinos Bardají, Antonio, U. de Zaragoza
Valera Pintor, Francisco, U. Carlos III
Vales Alonso, Javier, U. Pol de Cartagena
Vázquez Gallo, Enrique, U. Pol. de Madrid
Velasco Pérez, Juan R., U. Rey Juan Carlos
Verdú Pérez, Mª Jesús, U. de Valladolid
Vicario, Miguel, U. Aut. de Madrid
Vidal Catalá, José R., U. Pol. de Valencia
Vidal Fernández, Iván, U. Carlos III
Vidal Ferré, Rafael, U. Pol. de Catalunya
Villagrá González, Víctor, U. Pol. de Madrid
Vinyes Sanz, Joan, U. Pompeu Fabra
Vozmediano Torres, Juan M., U. de Sevilla
Yúfera Gómez, José M., U. Pol. de Catalunya

Agradecimientos

La organización agradece a las siguientes instituciones el patrocinio de las actividades de las Jornadas.

- Secretaría de Estado de Universidades e Investigación y Secretaría General de Política Científica y Tecnológica del Ministerio de Educación y Ciencia.
- Dirección Xeral de Universidades, Consellería de Educación e Ordenación Universitaria da Xunta de Galicia.
- Vicerrectorado de Investigación da Universidade de Vigo
- Centro Multimedia de Galicia, Dirección Xeral de Comunicación e Audiovisual, Consellería de Cultura, Comunicación Social e Turismo, Xunta de Galicia.
- Asociación de Técnicos de Informática

Asimismo, desea agradecer con especial estima el apoyo y la colaboración prestados *pro bono* por Cándido López García, Rebeca Díaz Redondo y Ana Fernández Vilas en beneficio de la organización general del evento.

Contenido

I. Ponencias

Redes inalámbricas, redes *ad hoc* y redes de sensores, I

Métodos docentes. Didáctica

Desarrollo de aplicaciones y servicios distribuidos. *E-learning*

Seguridad, control de acceso y prevención de ataques en red, I

Redes inalámbricas, redes *ad hoc* y redes de sensores, II

Distribución de contenidos. *Streaming*

Seguridad, control de acceso y prevención de ataques en red, II

Monitorización, medida y gestión de tráfico y de servicios

Arquitectura de redes de comunicaciones, I

Tecnología web

Arquitectura de redes de comunicaciones, II. Conmutación óptica

Descubrimiento de servicios, computación ubicua y domótica

Análisis de prestaciones y diseño de redes de comunicaciones

Ontologías y web semántica

II. Jornadas de seguimiento del Plan Nacional de I+D+I

Tecnología web, I

Tecnología web, II

Telemedicina

Tecnología educativa

Gestión y seguridad de red

Tecnologías de red, I

Tecnologías de red, II

Redes R-ALOHA DS-CDMA multicelulares con control de potencia rápido sobre canales Nakagami selectivos en frecuencia

Loren Carrasco, Guillem Femenias
Departament de Ciencies Matemàtiques i Informàtica. Universitat de les Illes Balears,
Ctra. Valldemossa km 7.5, 07122 Palma ESPAÑA,
Teléfono: 971 17 29 96 Fax: 971 17 30 03
E-mail: loren.carrasco@uib.es, guillem.femenias@uib.es.

Abstract *In this paper we investigate the performance of a multicell S-ALOHA DS-CDMA system with fast power control on a Nakagami frequency selective environment. We analyze how the throughput of the optimally stabilized system is affected by the channel conditions and some key system characteristics such as the power control error, the processing gain or the number of fingers in the RAKE receiver. By using an ideal retransmission probability that takes into account not only the number of users in backlog but also the DS-CDMA channel conditions we obtain the maximal achievable throughput of a S-ALOHA DS-CDMA system for a given DS-CDMA channel. These throughput figures can be used as an upperbound for currently used S-ALOHA DS-CDMA systems. The results obtained reflect that, if fast power control is used, the system throughput is very robust against varying channel conditions provided that the processing gain and the* $\frac{E_b}{N_0}$ *are above some threshold values.*

1. Introducción

Los protocolos ALOHA ranurados (R-ALOHA) sobre esquemas CDMA de secuencia directa (DS-CDMA) son utilizados en la actualidad para canales de acceso aleatorio para demandas o tráfico de datos a ráfagas dado que ofrecen la posibilidad de combinar las propiedades de R-ALOHA (p.e. simplicidad y acceso aleatorio) y DS-CDMA (p.e. multiplexación estadística) para conseguir una mayor eficiencia espectral [1].

A pesar de la gran cantidad de bibliografía disponible sobre protocolos R-ALOHA y sus variantes, la combinación de R-ALOHA y DS-CDMA ha recibido poca atención y existen todavía algunas cuestiones por resolver. En [2, 3] se proporciona un modelo muy útil para evaluar esquemas R-ALOHA DS-CDMA pero el análisis está restringido a un sistema unicelular no estabilizado sobre un canal AWGN. En [4] encontramos un estudio de la estabilidad de un sistema R-ALOHA de espectro ensanchado aunque la componente DS-CDMA no está correctamente modelada. En [5] se analiza un sistema multicelular sobre un canal Nakagami, aunque no se considera el funcionamiento del sistema una vez estabilizado y no se estudian las prestaciones en términos de throughput. Nuestro artículo difiere de la literatura previa en que realiza un modelado preciso de la capa física DS-CDMA y el canal de banda ancha asociado. De hecho al contrario que los análisis existentes (p.e.[6, 7, 8]) derivamos una expresión cerrada de la tasa media de bits erróneos (BER) considerando un control de potencia rápido con imperfecciones y un fading Nakagami selectivo en frecuencia.

Debido a la inherente inestabilidad de los esquemas ALOHA, las redes R-ALOHA DS-CDMA utilizan diferentes técnicas para estabilizar el sistema y permitir su correcto funcionamiento. Por tanto, estamos interesados fundamentalmente en las prestaciones de un sistema R-ALOHA DS-CDMA multicelular estabilizado y en como esas prestaciones se ven afectadas por el canal y la capa física DS-CDMA. Extendiendo el método presentado en [3], en este documento utilizaremos el mismo modelo de sistema R-ALOHA DS-CDMA que ya utilizamos en [9] y [10] pero en este caso investigaremos el comportamiento de un sistema con control de potencia rápido en lugar del control de potencia lento propuesto en [10]. Este modelo analítico nos permitirá obtener la expresión de la probabilidad de transmisión ideal que estabiliza el sistema de forma óptima, las figuras de throughput obtenidas en este sistema ideal son una cota superior para cualquiera de los métodos no ideales utilizados en la actualidad.

Este artículo se organiza de la forma siguiente: en la sección 2 se incluye una descripción global del sistema, a continuación en la sección 3 se describe el cálculo del throughput máximo y se determina la expresión de la tasa de error por paquete en función del BER. En la sección 4 se obtiene la expresión del BER y en la sección 5 se analizan las prestaciones del sistema. Finalmente en la sección 6 se incluyen las principales conclusiones de este estudio.

2. Descripción del modelo del sistema

En una red R-ALOHA DS-CDMA se asigna a cada estación base (BS) un conjunto de códigos de ensanchamiento (spreading) cuyas identidades se-

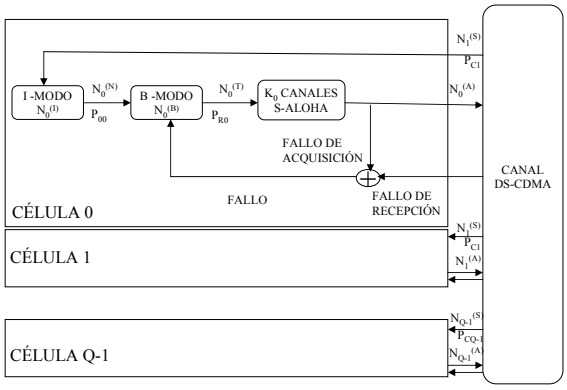


Figura 1: Modelo del sistema R-ALOHA DS-CDMA multicelular

rán difundidas por la BS a todos sus móviles (MS). Además el tiempo se divide en ranuras iguales en las que puede transmitirse exactamente un paquete. Cada vez que un móvil desea enviar un paquete a su BS lo ensancha seleccionando aleatoriamente uno de los códigos y lo transmite en la siguiente ranura. Si un conjunto de móviles transmiten en la misma ranura tendrán éxito:

- primero, si los MS utilizan diferentes códigos de ensanchamiento y los receptores de la BS son capaces de adquirirlos correctamente,
- y segundo, si esos paquetes correctamente adquiridos no presentan errores después del proceso de decodificación/detección del canal. La existencia de errores dependerá básicamente de la interferencia generada por múltiples transmisiones de MSs de la misma célula o células cercanas inherente a los esquemas DS-CDMA.

Los paquetes que no hayan tenido éxito deberán retransmitirse en una futura ranura.

El modelo del sistema queda completamente representado en la Fig.1 y sus parámetros descritos en la tabla 2. Consideramos una malla bidimensional de células hexagonales con una BS en el centro de cada célula. Utilizamos un sistema con N_{T0} terminales activos en la célula central (BS0) o célula-de-interés. Se asume que cada MS puede estar en uno de dos modos posibles al inicio de una ranura de tiempo: modo-I (Vacío/*Idle*, sin un paquete preparado para su transmisión) o modo-B (Espera/*Backlogged*, con un paquete preparado para su transmisión). Un MS que o bien 1) acaba de entrar en el sistema o 2) acaba de conseguir la transmisión exitosa de un paquete se dice que se encuentra en modo-I. Se asume que un MS en modo-I generará un paquete en el siguiente intervalo de tiempo y pasará a modo-B con probabilidad P_{00}. Por otro lado, los MSs en modo-B son aquellos que tienen paquetes esperando para su transmisión porque: 1) estaban en modo-I y han generado un paquete o 2) han tenido una transmisión sin éxito y están esperando una retransmisión. Se asume que un MS en modo-B transmitirá un paquete en el siguiente intervalo con probabilidad P_{R0}. También asumimos que un terminal en modo-B no genera nuevos paquetes. Se dice que un terminal en modo-B que transmite/retransmite un paquete ha adquirido uno de los K_0 códigos disponibles en la BS0 si no hay otros MSs de esa célula en modo-B con una transmisión en curso que en el mismo intervalo hayan seleccionado el mismo código.

Las transmisiones en BS0 se verán afectadas por una cierta tasa de paquetes transmitidos con éxito (P_{C0}) que depende de las características del canal y el nivel de interferencia. Para $N_{0,k}^{(T)}$ transmisiones simultáneas en la ranura k en la célula $q = 0$, hay $N_{0,k}^{(T)} - 1$ interferentes intracelulares y $M_0 T$ interferentes intercelulares. Asumiendo que las células están sincronizadas a nivel de ranura, $M_0 T$ es la suma de los MSs transmitiendo en la misma ranura pero no conectados a BS0.

El canal radio utilizado para calcular P_{Cq} está afectado por fluctuaciones a corto y largo término. Las variaciones a largo término son debidas a las pérdidas con la distancia con índice μ y a los desvanecimientos lentos *shadowing*, modelados como una variable aleatoria log-normal de la forma $10^{\xi/10}$ donde se asume que ξ sigue una distribución Gaussiana de media cero y varianza σ_{sh}^2 . Asumimos también que la multipropagación supone la recepción de L caminos con una distribución Nakagami que serán combinados de forma óptima por receptores RAKE con L etapas. Cada BS transmite una señal piloto y las MSs se conectan a la BS cuyo piloto reciben con mayor potencia media. Dado que una MS puede no ser capaz de medir el piloto de todas las BSs del sistema, la selección se considerará limitada a las Q_C BSs más cercanas. Esta limitación puede justificarse también por el uso por parte de los MSs de listas de células vecinas consideradas como posibles candidatas en el proceso de handover [7][11]. Definimos como S_0, con área A_0, la región del plano que contiene todos los puntos que tienen la base central (BS0) dentro de su conjunto de Q_C BSs más próximas, y definimos como S_1, con área A_1, la región que contiene todos los puntos que no tienen la BS0 dentro de las Q_C BSs más cercanas. El tamaño y forma de estas regiones es una función de Q_C y pueden encontrarse ejemplos en [7][11]. Además asumimos que el sistema cuenta con un control de potencia rápido gracias a un canal de señalización lento existente entre cualquier MS y su BS. Consideraremos que existe un error en el control de potencia (PCE) que se modelará como una variable aleatoria lognormal.

3. Análisis del sistema

Suponemos que la célula-de-interés ($q = 0$) se encuentra en el estado B, lo que significa que el número de MSs en espera es $N_{0,k}^{(B)} = B$. Supongamos ahora que $S_0^{in}(B)$ y $S_0^{out}(B)$ son, respectivamente, el

Param	Descripción
Q	N° células sistema
N_q	N° MSs célula q
K	Num. pares receptor-código cel. q
$N_{q,k}^{(B)}$	Num. MSs modo-B inicio ranura k célula q
$N_{q,k}^{(I)}$	Num. MSs modo-I inicio ranura k célula q
$N_{q,k}^{(T)}$	Num. Total MSs transmitiendo paquetes ranura k célula q
$N_{q,k}^{(N)}$	Num. MSs modo-I generando paquete ranura k célula q
$N_{q,k}^{(A)}$	Num. paquetes adquiridos ranura k cél. q
$N_{q,k}^{(S)}$	Num. paquetes correctamente recibidos ranura k célula q
P_{0q}	Prob. MS modo-I genere un paquete cél. q
P_{Rq}	Prob. MS modo-B transmita un paquete cél. q
M_{qT}	Num. interferentes intercelulares ranura k célula q
P_{Cq}	Prob. media transmisión con éxito de paquete célula q

Cuadro 1: Descripción de los parámetros del sistema

número medio de paquetes netos que fluyen hacia el sistema ($N_{0,k}^{(N)}$ medio) y el número medio de paquetes que fluyen fuera del sistema ($N_{0,k}^{(S)}$ medio) en una ranura temporal. Si $S_0^{in}(B) > S_0^{out}(B)$, entonces el sistema tiende a desplazarse hacia un estado con un B mayor ($> B$) en cambio si $S_0^{in}(B) < S_0^{out}(B)$ el sistema tenderá a desplazarse hacia un estado con un B menor($< B$). Si $S_0^{in}(B) = S_0^{out}(B)$, entonces el estado con B usuarios en espera es un estado de equilibrio. Un estado de equilibrio puede ser tanto estable como inestable. Si el número de estados de equilibrio es uno, se dice que el sistema es estable. En cualquier otro caso se dice que es inestable. Hay que subrayar que cuando el sistema alcanza un punto estable B_e el throughput medio coincide con $S_0^{out}(B_e)$.

El número medio de paquetes fluyendo fuera del sistema $S_0^{out}(B)$ es [9]

$$S_0^{out}(B) = \sum_{n=0}^{B} \sum_{s=0}^{\min(K_0,n)} s\binom{B}{n} P_{R0}^n (1-P_{R0})^{(B-n)} \times Pr\{N_{0,k}^{(S)}=s|N_{0,k}^{(T)}=n\}, \quad (1)$$

donde

$$Pr\{N_{0,k}^{(S)}=s|N_{0,k}^{(T)}=n\} = \sum_{a=0}^{\min(K_0,n)} \binom{a}{s} P_{C0}(n)^s \times [1\text{-}P_{C0}(n)]^{a-s} \frac{\binom{K_0}{a}\binom{n}{a} a! T_{n\text{-}a,K_0\text{-}a}}{K_0^n}, \quad (2)$$

si $0 \le s \le a$ y 0 en cualquier otro caso, $P_{C0}(n)$ es la tasa de paquetes transmitidos con éxito para n transmisiones simultáneas en la célula de interés ($N_0^{(T)} = n$) y

$$T_{x,y} = y^x - \left[\sum_{i=1}^{\min(x,y)} \binom{y}{i}\binom{x}{i} i! T_{x-i,y-i} \right]. \quad (3)$$

Derivando (1) con respecto a P_{R0} en [9] obtenemos

$$\frac{\partial S_0^{out}(B)}{\partial P_{R0}} = \sum_{n=0}^{N_{T0}} \sum_{s=0}^{\min(K_0,n)} s\binom{B}{n} \Big[nP_{R0}^{n\text{-}1}(1\text{-}P_{R0})^{(B\text{-}n)} - P_{R0}^n (1\text{-}P_{R0})^{(B\text{-}n\text{-}1)}(B\text{-}n) \Big] Pr\{N_{0,k}^{(S)}=s|N_{0,k}^{(T)}=n\}, \quad (4)$$

y a partir de los ceros de esta derivada conseguimos una probabilidad de retransmisión adaptativa P_{R0}^* que garantiza la estabilidad y maximiza el throughput de BS0.

Para investigar las prestaciones del sistema utilizaremos P_{R0}^* con el fin de estabilizar el sistema en BS0 y entonces podremos determinar el throughput resultante en BS0 utilizando (1). Sin embargo la determinación del throughput requiere el cálculo de la tasa de paquetes que se transmiten con éxito para cualquier número de transmisiones simultáneas en la célula $P_{C0}(n)$ con $n \in \{1, 2, ..., N_{T0}\}$. La obtención de $P_{C0}(n)$ requiere calcular la tasa media de bits erróneos $P_b(N_0^{(T)} = n)$ para todos los valores posibles de n. Entonces, si asumimos que la resistencia a los desvanecimientos rápidos de la tecnología DS-CDMA y el entrelazado existente provocarán que los errores sean independientes y estén idénticamente distribuidos dentro de un paquete, y si además suponemos que se utiliza un código bloque t corrector, la probabilidad de éxito de un paquete $P_{C0}(n)$ para paquetes con E bits está acotada por [12]

$$P_{C0}(n) \le \sum_{i=0}^{t} \binom{E}{i} P_b(n)^i (1 - P_b(n))^{E-i}. \quad (5)$$

4. Análisis de la probabilidad de error

4.1. Modelos del canal y del transmisor-receptor

Utilizamos un canal selectivo en frecuencia con una respuesta impulsional

$$h_{q0,k}(\tau, t) = \frac{\sqrt{G_c}}{d_{k0}^{\mu/2}(x,y) 10^{\xi_{k0}/20}} \times \sum_{l=0}^{L_{q0,k}-1} \alpha_{q0,kl}(t) e^{j\psi_{q0,kl}(t)} \delta(\tau\text{-}lT_c) \quad (6)$$

donde G_c es una constante, μ es el índice de pérdidas con la distancia, $\xi_{k,0} \sim N(0, \sigma^2_{sh_{k,0}})$ corresponde al desvanecimiento lognormal, T_c se corresponde

con el período de chip y por tanto $f_s = 1/T_c$ es el ancho de banda de la señal transmitida real paso banda, $L_{q0,k} = \lfloor f_s/(\Delta f)_{C_{q0,k}} \rfloor$ es el número de caminos que pueden resolverse separadamente en ese canal ya que $(\Delta f)_{C_{q0,k}}$ es el ancho de banda de coherencia del canal. Si asumimos que la duración del símbolo es mucho menor que el tiempo de coherencia del canal, entonces $\alpha_{q0,kl}(t)e^{j\psi_{q0,kl}(t)} = \alpha_{q0,kl}e^{j\psi_{q0,kl}}$, donde $\{\psi_{q0,kl}\}$ son variables aleatorias independientes uniformemente distribuidas entre $[0, 2\pi)$ y $\{\alpha_{q0,kl}\}$ son variables aleatorias independientes Nakagami [13] con parámetros, $\Omega_{q0,kl} = \mathrm{E}\{\alpha^2_{q0,kl}\}$ y $m_{q0,kl} = \Omega^2_{q0,kl}/\mathrm{E}\{(\alpha_{q0,kl} - \Omega_{q0,kl})^2\}$ ($m_{q0,kl}$ es el parámetro que determina la severidad de los desvanecimientos rápidos en el camino l entre el MS k de la BSq y la BS0 y las $\Omega_{q0,kl}$ están relacionadas con el perfil de intensidad de la multipropagación (MIP). Contrariamente a [7, 8] no impondremos ninguna restricción sobre la forma del MIP.

La señal equivalente en banda base transmitida por un MS k en una célula q (enlace ascendente) puede expresarse como

$$\widetilde{s}_k(t) = \sqrt{\frac{2S\lambda_k}{G_c\Phi_{k\hat{q}}}}\, d^{\frac{\mu}{2}}_{k\hat{q}}(x,y)10^{\frac{\xi_{k\hat{q}}}{20}} D_k(t\text{-}\tau_k)c_k(t\text{-}\tau_k)e^{j\phi_k} \tag{7}$$

donde S representa la potencia de la señal recibida con control de potencia ideal, $D_k(t)$ es la forma de onda de los datos codificados del MS k, $c_k(t)$ es la forma de onda de la correspondiente secuencia de código, τ_k indica que cada MS tiene una temporalización independiente debido a que el sistema funciona con transmisión asíncrona, λ_k modela la amplitud del error en el control de potencia (PCE), que asumimos sigue una distribución lognormal [6, 14, 7], y que por tanto puede escribirse como $\lambda_k = 10^{x_k/10}$ donde x_k es una variable aleatoria gaussiana de media nula y desviación típica σ_{e_k}, $\phi_k = \theta_k - \omega_c\tau_k$, donde θ_k es la fase de la portadora y ω_c la frecuencia de dicha portadora. Finalmente el término $d^{\mu/2}_{k\hat{q}}(x,y)10^{\xi_{k\hat{q}}/20}$ corresponde a la compensación por parte del control de potencia del MS de las pérdidas con la distancia y el desvanecimiento lognormal hacia su propia BS (definida como $\hat{q}$) y de la misma forma el término $\Phi_{k\hat{q}} = \sum_{l=0}^{L_{\hat{q}\hat{q},k}-1} \alpha^2_{\hat{q}\hat{q},kl}$ corresponde a la compensación de los desvanecimientos rápidos.

Para alcanzar la BS0, la señal transmitida se ve afectada por el canal descrito anteriormente y entonces la señal recibida en BS0 puede expresarse como

$$\widetilde{r}(t) = \sum_{\forall k}\sqrt{\frac{2S\lambda_k\Upsilon_k(x,y)}{\Phi_{k\hat{q}}}}e^{j\phi_k}\sum_{l=0}^{L_{q0,k}-1}\alpha_{q0,kl}e^{j(\psi_{q0,kl})} \times D_k(t - lT_c - \tau_k)c_k(t - lT_c - \tau_k) + \widetilde{n}(t), \tag{8}$$

donde $\widetilde{n}(t)$ es el ruido blanco gaussiano aditivo complejo de media cero y densidad espectral de potencia unilateral η_0 y

$$\Upsilon_k(x,y) \triangleq \begin{cases} 1 & ,k \in \mathcal{S}_{\mathrm{BS0}} \\ \dfrac{\min\limits_{\substack{q\in\Theta_k \\ q\neq 0}}\left\{d^{\mu}_{kq}(x,y)10^{\xi_{kq}/10}\right\}}{d^{\mu}_{k0}(x,y)10^{\xi_{k0}/10}} & ,k \in \mathcal{S}_{\overline{\mathrm{BS0}}} \\ \dfrac{\min\limits_{q\in\Theta_k}\left\{d^{\mu}_{kq}(x,y)10^{\xi_{kq}/10}\right\}}{d^{\mu}_{k0}(x,y)10^{\xi_{k0}/10}} & ,k \in \mathcal{S}_1 \end{cases} \tag{9}$$

con $\mathcal{S}_{\mathrm{BS0}}$ representando el conjunto de MSs situados en $\mathcal{S}_0$ que están conectados a la BS0 y $\mathcal{S}_{\overline{\mathrm{BS0}}}$ representando el conjunto de MSs en $\mathcal{S}_0$ no conectados a BS0.

Si asumimos una estimación perfecta del canal, la salida del receptor de correlación para el usuario deseado (q=0, k=1), obtenido sin pérdida de generalidad en el instante de muestreo T_b es

$$r(T_b) = \mathrm{Re}\left\{\sum_{l=0}^{L_{00,1}-1}\alpha_{00,1l}e^{j(\psi_{00,1l}+\phi_1)} \times \int_{lT_c+\tau_1}^{T_b+lT_c+\tau_1}\widetilde{r}(t)c_1(t - lT_c - \tau_1)dt\right\} \tag{10}$$

Esta expresión puede descomponerse en los siguientes términos

$$r(T_b) = r_u(T_b)+r_{mp}(T_b)+r_{ma}(T_b)+r_{mc}(T_b)+r_{th}(T_b) \tag{11}$$

donde $r_u(T_b) = \sqrt{2S\lambda_1}T_b$ corresponde a la aportación de la señal útil, $r_{mp}(T_b)$ representa la autointerferencia debida a la multipropagación, $r_{ma}(T_b)$ y $r_{mc}(T_b)$ corresponden a la interferencia de otras MSs en la misma célula y células vecinas respectivamente, y finalmente $r_{th}(T_b)$ es la variable aleatoria Gaussiana debida al proceso AWGN ($Re\{r_{th}(T_b)\} = N(0, T_b\eta_0\sum_{l=1}^{L_{00,1}}\alpha^2_{00,1l})$). Dado que $(n-1) + \sum_{q=1}^{Q} M_q >> 1$, $r_{mp}(T_b)$ es mucho más pequeña que $r_{ma}(T_b)$ y $r_{mc}(T_b)$ y por tanto se despreciara su aportación a partir de ahora.

4.2. Hipótesis Gaussiana

Definiendo $r_{in}(T_b) = r_{mc}(T_b)+r_{ma}(T_b)$ y teniendo en cuenta que cada MS sufre unos desvanecimientos y PCE independientes de los demás MSs y que su localización es también independiente de la de los demás, entonces $r_{in}(T_b)$ es una suma de variables aleatorias independientes que aplicando el teorema central del límite asumiremos que tiene una distribución Gaussiana. La utilización de la hipótesis gaussiana en los cálculos de la tasa de error es muy común [15],[8], ya que se ha determinado que se trata de una aproximación bastante exacta. Por tanto $r_{in}(T)$ es asintóticamente Gaussiana condicionada al PCE del MS-de-interés. Asumiendo

secuencias de ensanchamiento largas, un MIP normalizado ($\sum_{n=0}^{L_0,k-1} \mathrm{E}\{\alpha_{q0,kn}^2\} = 1$) para todos los MSs, asumiendo también que $E\left\{\frac{1}{\Phi_{kq}}\right\} = E\left\{\frac{1}{\Phi}\right\}$ y $E\{\lambda_k\} = e^{\frac{1}{2}\mathcal{B}^2\sigma_e^2}$, donde $\mathcal{B} \triangleq \frac{\ln 10}{10}$, la varianza condicional de la variable $r_{in}(T)$ de media cero es

$$\mathrm{Var}[r_{in}(T_b)] = \frac{4G_pT_c^2S}{3}\varepsilon e^{\frac{1}{2}\mathcal{B}^2\sigma_e^2}E\left\{\frac{1}{\Phi}\right\} \times \sum_{l=0}^{L_{00,1}-1} \alpha_{00,1l}^2\left((n-1) + \varpi_{\mathcal{S}_{\overline{\mathrm{BS0}}}}M_{\overline{\mathrm{BS0}}} + \varpi_{\mathcal{S}_1}M_{S_1}\right) \tag{12}$$

donde Gp es la ganancia de procesado, $n = N_0^{(T)}$, $M_{\overline{\mathrm{BS0}}}$ es el número de MSs en S_0 que no están conectadas a $BS0$, M_{S_1} es el número de MSs en S_1 y $\varpi_{\mathcal{S}_\chi} \triangleq E\{\Upsilon_k(x,y)\}_{k\in\mathcal{S}_\chi}$ denota la interferencia media producida por un MS situado en $\mathcal{S}_\chi$

$$\varpi_{\mathcal{S}_\chi} = \begin{cases} \frac{1}{A_0}\iint\limits_{S_0} E\left\{e^{\zeta_{\mathcal{S}_0}^{(k)}} | \zeta_{\mathcal{S}_0}^{(k)} < 0\right\} dA_0 & , \mathcal{S}_\chi{=}\mathcal{S}_{\overline{\mathrm{BS0}}} \\ \frac{1}{A_1}\iint\limits_{\mathcal{S}_1} E\left\{e^{\zeta_{\mathcal{S}_1}^{(k)}}\right\} dA_1 & , \mathcal{S}_\chi{=}\mathcal{S}_1 \end{cases} \tag{13}$$

donde

$$\zeta_{\mathcal{S}_0}^{(k)} \triangleq -\mathcal{B}\xi_{k0} + \min_{\substack{q\in\Theta_k \\ q\neq 0}}\left\{\mu \ln\frac{d_{kq}}{d_{k0}} + \mathcal{B}\xi_{kq}\right\}, \tag{14}$$

$$\zeta_{\mathcal{S}_1}^{(k)} \triangleq -\mathcal{B}\xi_{k0} + \min_{q\in\Theta_k}\left\{\mu \ln\frac{d_{kq}}{d_{k0}} + \mathcal{B}\xi_{kq}\right\}. \tag{15}$$

El SNIR a la salida del receptor puede escribirse como $\gamma_b = \Psi_n\lambda_1$, donde Ψ_n es una constante dada por

$$\Psi_n = \frac{E\left\{\frac{1}{\Phi}\right\}^{-1}}{\frac{e^{\frac{1}{2}\mathcal{B}^2\sigma_e^2}}{3G_p}\left(n\text{-}1 + \varpi_{\mathcal{S}_{\overline{\mathrm{BS0}}}}M_{\overline{\mathrm{BS0}}} + \varpi_{\mathcal{S}_1}M_{S_1}\right) + \frac{N_0}{E_b}}. \tag{16}$$

donde $E_b = 2ST_bE\left\{\frac{1}{\Phi}\right\}$ es la energía media recibida por bit. La variable aleatoria Φ es una suma de variables gamma, podemos utilizar por tanto el teorema de Moschopoulos [16] para encontrar su función de densidad de probabilidad

$$p_\Phi(\zeta) = \sum_{k=0}^{\infty} \frac{\prod_{n=1}^{L_{00,1}}\left(\frac{\beta_1}{\beta_n}\right)^{m_n}\delta_k e^{-\zeta/\beta_1}\zeta^{\mathcal{S}_k-1}}{\Gamma(\mathcal{S}_k)(\beta_1)^{\mathcal{S}_k}} \tag{17}$$

donde $\beta_n = \Omega_{\hat{q}\hat{q},kn}/\alpha_{\hat{q}\hat{q},kn}$, $\beta_1 = \min_n\{\beta_n\}$, $\mathcal{S}_k = \sum_{n=1}^{L_{00,1}} m_n + k$ y los coeficientes δ_k pueden obtenerse recursivamente con la fórmula

$$\begin{cases} \delta_0 & = 1 \\ \delta_{k+1} & = \frac{1}{k+1}\sum_{i=1}^{k+1}\left[\sum_{j=1}^{N} m_j\left(1-\frac{\beta_1}{\beta_j}\right)^i\right] \end{cases} \tag{18}$$

Moschopoulos [16] proporciona una prueba rigurosa de la convergencia uniforme de (17) y una cota del error de truncado. Utilizando (17) obtenemos

$$\mathrm{E}\left[\frac{1}{\Phi}\right] = \int_0^\infty \frac{p_\Phi(\Phi)}{\Phi}d\Phi = \sum_{k=0}^{\infty}\delta_k\beta_1^{-1}\prod_{n=1}^{L}\frac{\beta_1}{\beta_n}^{m_n}\frac{\Gamma(\mathcal{S}_k\text{-}1)}{\Gamma(\mathcal{S}_k)} \tag{19}$$

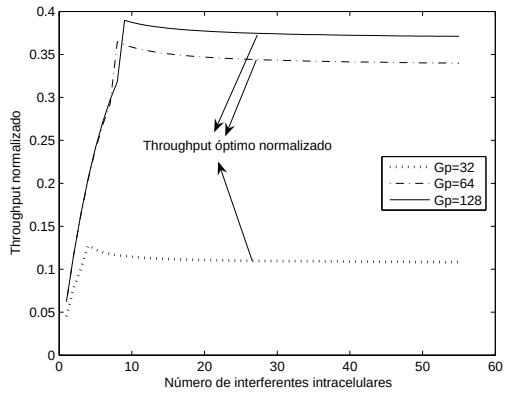


Figura 2: Throughput normalizado vs número de interferentes intracelulares en función de Gp

4.3. Cálculo BER

Es bien conocido [17] que si se utiliza modulación BPSK coherente en presencia de AWGN, la probabilidad de error condicionada al SNIR instantáneo puede expresarse como

$$P_b(n|\lambda_1) = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\Psi_n\lambda_1}\right) \tag{20}$$

Considerando que λ_1 es una variable aleatoria lognormal, podemos utilizar la expansión de diferencias propuesta en [18] para obtener una aproximación precisa al promedio de la probabilidad de error por bit

$$P_b(n) = \frac{2}{3}P_b(n|1) + \frac{1}{6}P_b(n|\sqrt{3\sigma_e^2}) + \frac{1}{6}P_b(n|-\sqrt{3\sigma_e^2}) \tag{21}$$

5. Resultados Numéricos

En esta sección se estudiará el impacto de diferentes parámetros del canal y la capa física DS-CDMA sobre las prestaciones del sistema. Los parámetros considerados por defecto son: un mallado celular con $Q = 91$ células hexagonales, un canal con un MIP exponencial del tipo $\Omega_l = \Omega_1 e^{-0{,}5(l-1)}$, donde Ω_l es la potencia instantánea del camino l, un número de caminos $L = 3$, un parámetro Nakagami $m = 1{,}5$, una ganancia de procesado $G_p = 256$, un exponente de pérdidas con la distancia $\mu = 4$, una desviación del desvanecimiento lognormal $\sigma_{sh} = 6dB$, una desviación en el control

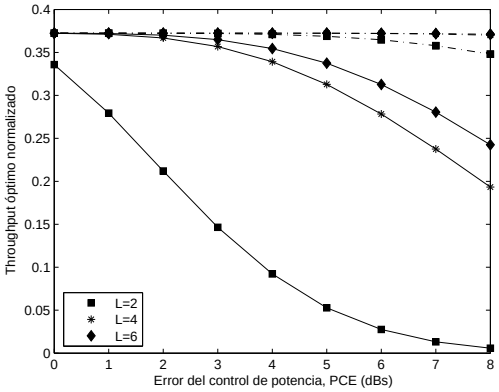


Figura 3: Throughput óptimo normalizado vs el número de caminos (L) en función del PCE

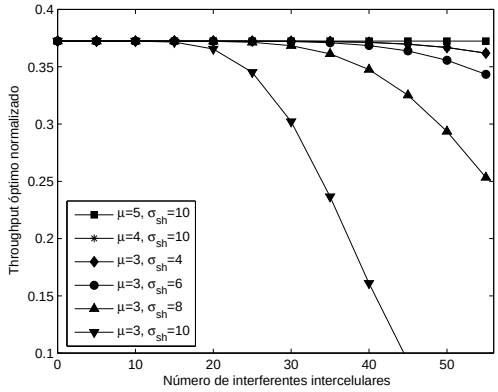


Figura 4: Throughput óptimo normalizado vs el número de interferentes intercelulares en función de σ_{sh} y μ

de potencia $\sigma_e = 1dB$, un conjunto de BSs seleccionables $Q_c = 4$ y una longitud de paquete $E = 1024$ bits.

La Fig.2 muestra las curvas de throughput normalizado (número de paquetes por código CDMA y ranura) en función de la carga de la célula para diferentes ganancias de procesado de un sistema estabilizado utilizando (4), puede verse que una vez se ha alcanzado el punto de saturación el valor del throughput de todas las curvas se mantiene prácticamente constante para cualquier carga en la célula. Todas las demás figuras representan la variación sufrida por este throughput óptimo normalizado vs diferentes parámetros. La fig.3 representa el efecto del PCE sobre el throughput del sistema para diferentes valores L considerando un número fijo de interferentes intercelulares ($M_T = 20$), la figura muestra dos grupos de curvas: las curvas con líneas de puntos corresponden a las condiciones por defecto del canal mientras que las curvas con líneas continuas corresponden a un canal peor ($m = 0{,}75$, $\mu = 3$, $\sigma_{sh} = 8$). Es evidente que la degradación debida al PCE tiene lugar solamente para malas condiciones del canal como cuando hay pocos caminos disponibles (curva con L=2) en el receptor. La Fig.4 muestra el efecto del índice de pérdidas con la distancia μ y el desvanecimiento lognormal (σ_{sh}) sobre el throughput en función del número de interferentes intercelulares. En este caso se muestra una degradación del throughput sólo cuando $\mu \leq 3$ y $\sigma_{sh} \geq 8$. Para valores de μ bajos las interferencias intercelulares se atenúan menos en su camino hacia la célula-de-interés y este efecto combinado con un desvanecimiento lognormal importante provoca un incremento del nivel de interferencia producida por cada MS interferente y por tanto una degradación en las prestaciones del sistema cuando el número de interferentes aumenta. El efecto de los desvanecimientos lognormales se debe a que un desvanecimiento lognormal más elevado implica una mayor probabilidad de que los MSs no estén conectados al BS que reciben con una atenuación media menor, incrementando de esta forma la potencia emitida

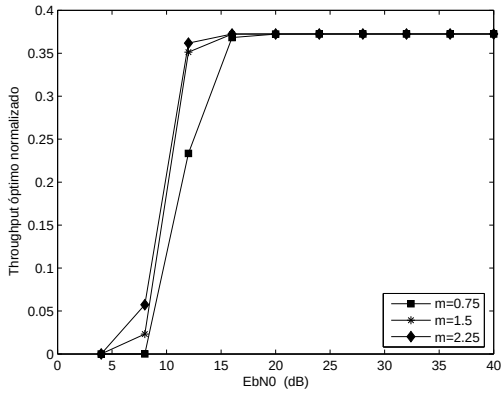


Figura 5: Throughput óptimo normalizado vs el $\frac{E_b}{N_0}$ en función de m

por esos MSs. El efecto del valor de la $\frac{E_b}{N_0}$ combinada con el parámetro Nakagami m está representado en la figura 5. Esta figura muestra como existe un umbral de $\frac{E_b}{N_0}$ que permite la operación correcta del sistema y que este umbral varia con las condiciones del canal, por ejemplo en la gráfica se muestra como se necesita un $\frac{E_b}{N_0} \geq 15dB$ en el peor de los canales considerados (m=0.75, peor que un canal Rayleigh) mientras que un $\frac{E_b}{N_0} \geq 10dB$ es suficiente para canales mejores (m=1.5-2.25)

Finalmente la Fig.6 muestra el efecto de la ganancia de procesado G_p y su habilidad para mitigar el impacto de las interferencias. Esta gráfica muestra también que al igual que con el $\frac{E_b}{N_0}$ se requiere una ganancia de procesado mínima $G_p > 128$ para permitir la operación correcta del sistema.

6. Conclusiones

En este artículo hemos investigado las prestaciones de un sistema R-ALOHA DS-CDMA multicelular con control de potencia rápido sobre un canal selectivo en frecuencia. Hemos analizado también cómo el throughput del sistema estabilizado

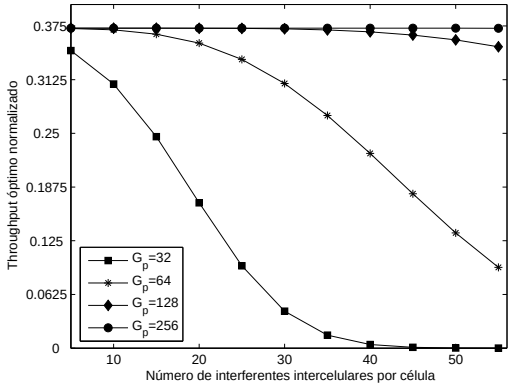


Figura 6: Throughput óptimo normalizado vs el número de interferentes intercelulares en función de G_p

de forma óptima se ve afectado por las condiciones del canal y algunas características clave del sistema como la ganancia de procesado o el número de etapas del receptor RAKE. Utilizando una probabilidad de retransmisión ideal que tiene en cuenta no sólo el número de usuarios en espera sino también las condiciones del canal DS-CDMA obtenemos el throughput máximo posible del sistema sobre un determinado canal. Estos resultados pueden utilizarse como cota superior para sistemas R-ALOHA DS-CDMA utilizados en la actualidad.

Los resultados obtenidos reflejan que si se utiliza un control de potencia rápido el throughput del sistema se ve poco afectado por variaciones del canal DS-CDMA excepto cuando la degradación es muy significativa, debida por ejemplo a una ganancia de procesado insuficiente $Gp < 128$, a una figura de ruido que no alcanza un determinado umbral $\frac{E_b}{N_0} < 15dB$, o si se produce una combinación de algunas de las siguientes condiciones: un PCE muy significativo ($PCE \geq 2dBs$), un gran número de interferentes intercelulares ($M_{qT} > 20$ para cualquier célula interferente q), un nivel importante de desvanecimiento lognormal ($\sigma_{sh} \geq 8dBs$) y desvanecimientos rápidos ($m \leq 1$), un bajo índice de pérdidas con la distancia ($\mu \leq 3$) o un pequeño número de caminos disponibles en el receptor ($L < 3$).

Agradecimientos

Este trabajo ha sido financiado en parte por el Ministerio de Ciencia y Tecnología y FEDER (Fondo Europeo de Desarrollo Regional) dentro del proyecto (TIC2001-0287).

Referencias

[1] Raychaudhuri, D.: Performance analysis of random access packet-switched code division multiple access systems. IEEE Trans.Commun. **29** (1981) 895–901

[2] Z.Liu, Zarki, M.E.: Performance analysis of DS-CDMA with slotted ALOHA random access for packet PCNs. Wireless Networks 1,J.C Baltzer AG, Science Publishers (1996) 1–16

[3] Dastangoo, S., Vojcic, B., Daigle, J.: Performance analysis of multi-code spread slotted ALOHA (MCSSA) system. Wireless Networks 1 (1995) 1–16

[4] de Graaf, W., Lehnert, J.: Performance comparison of a slotted ALOHA DS/SSMA network and multichannel narrow-band slotted ALOHA network. IEEE Trans.Commun. **46** (1998) 544–552

[5] Cooper, W., Zeidler, J., McLaughlin, S.: Performance analysis of slotted random access channels for W-CDMA systems in nakagami fading channels. IEEE Trans. Veh. Technol. (2002) 411–424

[6] Corazza, G., Maio, G.D., Vatalaro, F.: CDMA cellular system performance with fading, shadowing and imperfect power control. IEEE Trans. Veh. Technol. **47** (1998) 450–459

[7] Romero-Jerez, J., Tellez-Labao, C.: Effect of power control imperfections on the reverse link of cellular CDMA networks under multipath fading. IEEE Trans. Veh. Technol. **53** (2004) 61–71

[8] Kong, N., Milstein, L.: Error probability of multicell CDMA over frequency selective fading channels with power control error. IEEE Trans.Commun. **47** (1999) 608–617

[9] Carrasco, L., Femenias, G.: Stability analysis of slotted ALOHA DS-CDMA communication networks. Proceed. ICWN02 (2002) 176–182

[10] Carrasco, L., Femenias, G.: Multicell S-ALOHA DS-CDMA networks over frequency selective nakagami channels with open loop power control error. Proceed. PIMRC04 (2004)

[11] Femenias, G., Carrasco, L.: Effect of slow power control error on the reverse link of STBC DS-CDMA in a cellular system with nakagami frequency-selective MIMO fading. (Submitted to IEEE Trans. Veh. Technol.)

[12] Lin, S., Costello, D.: Error Control Coding: Fundamentals and Applications. Prentice Hall Series in computer applications in electrical engineering (1983)

[13] Nakagami, M.: The m-distribution—a general formula of intensity distribution of rapid fading. In Hoffman, W.C., ed.: Statistical study of radio wave propagation, Pergamon Press (1960) 3–36

[14] Abrardo, A., Sennati, D.: On the analytical evaluation of closed-loop-power-control error statistics in DS-CDMA cellular systems. IEEE Trans. Veh. Technol. **49** (2000) 2071–2080

[15] Kavehrad, M., McLane, P.: Performance of low complexity channel coding and diversity for spread spectrum in indoor, wireless communications. AT&T Technical Journal **64** (1985) 1927–1964

[16] Moschopoulos, P.: The distribution of the sum of independent gamma random variables. Ann. Inst. Statist. Math. **37** (1985) 541–544

[17] Proakis, J.G.: Digital communications. 2nd edn. McGraw-Hill, Singapore (1989)

[18] Holtzman, J.M.: A simple, accurate method to calculate spread-spectrum multiple access error probabilities. IEEE Trans. Commun **40** (1992) 461–464

Asignación Eficiente de Recursos para los Servicios de Broadcast y Punto a Punto en el Protocolo ADHOC MAC

José Ramón Gállego, Ángela Hernández-Solana, María Canales, Antonio Valdovinos
Departamento de Ingeniería Electrónica y Comunicaciones, Universidad de Zaragoza.
María de Luna 1, 50018 Zaragoza
E-mail: jrgalleg@unizar.es

Luca Campelli, Matteo Cesana, Antonio Capone, Flaminio Borgonovo
Dipartimento di Elettronica e Informazione, Politecnico di Milano
Piazza L. Da Vinci 32, 20133 Milán, Italia
E-mail: campelli@elet.polimi.it

Abstract. *An effective Medium Access Control for communications in wireless Ad hoc networks should be able to support both broadcast and point-to-point communications paradigms. The ADHOC MAC protocol, recently proposed within the European Commission funded CarTALK2000 project, seems to match these requirements. As a matter of fact, it allows the exchange of connectivity information among wireless terminals which can be usefully exploited to devise both broadcast and point-to-point services. In this paper we evaluate through simulation the efficiency of the protocol in a mixed traffic scenario where broadcast and point-to-point communications coexist. An adaptive bandwidth allocation strategy is proposed to share the resources between both services in a dynamic situation. The capability of the protocol to establish parallel point-to-point data communications and the corresponding improvement in the point-to-point efficiency is also evaluated.*

1 Introducción

El medio de transmisión en entornos inalámbricos tiene que ser compartido por definición. Además, los recursos radio son a menudo limitados frente al número de usuarios que intentan acceder a los mismos, lo que hace que la capacidad de cualquier red inalámbrica esté altamente determinada por la capacidad de los mecanismos de control de acceso al medio para gestionar dicho proceso de acceso y conseguir un alto reuso de los recursos [1].

ADHOC MAC [2] es un protocolo de acceso al medio propuesto recientemente dentro del proyecto europeo CarTalk 2000 [3] para proporcionar conectividad en redes ad hoc inter-vehiculares [4]. ADHOC MAC funciona sobre una capa física síncrona e implementa una técnica de acceso completamente distribuida capaz de establecer dinámicamente un canal de broadcast fiable (Basic broadcast Channel: BCH) para cada terminal activo. Cada BCH contiene información de señalización que proporciona una distribución de la información de conectividad de la red rápida y fiable a todos los terminales. Esta información proporciona una base sólida para la implementación de servicios de datos punto a punto, explotando las transmisiones paralelas, y facilita la gestión de distintos requerimientos de QoS para estos servicios mediante el uso de prioridades.

En [5] y [6] se han estudiado las prestaciones de los servicios de broadcast en ADHOC MAC en un escenario estático y con movilidad respectivamente. En este artículo, evaluamos mediante simulación la eficiencia del protocolo en un escenario de tráfico mixto donde coexisten las comunicaciones broadcast y punto a punto. Se propone una estrategia adaptativa de asignación de ancho de banda para compartir los recursos entre ambos servicios en una situación dinámica. El objetivo de dicha propuesta es garantizar los requerimientos de acceso para el BCH maximizando la capacidad para comunicaciones de datos adicionales. También es evaluada la capacidad del protocolo para establecer comunicaciones punto a punto paralelas y la correspondiente mejora en la eficiencia que esto implica. El resto del artículo está organizado del modo siguiente: en la sección 2 describimos brevemente las bases el protocolo ADHOC MAC y las estrategias de asignación de ancho de banda propuestas para el servicio básico de broadcast y los servicios punto a punto. En la sección 3, tanto las estrategias de gestión de recursos como la eficiencia del servicio punto a punto son evaluadas mediante simulación. Finalmente, en la sección 4 se presentan las conclusiones principales.

2 El Protocolo ADHOC MAC

2.1 Modo de operación básico para BCH y comunicaciones punto a punto

ADHOC MAC está basado en una estructura de slots temporales, en la que los slots están agrupados en tramas virtuales de longitud *N* y que, en principio, no requiere alineación de trama. En el BCH, cada terminal envía en broadcast la información del estado del canal que él percibe. El BCH contiene un campo de control, Frame Information (FI), que es un vector de *N* elementos que especifica el estado de los *N* slots que preceden a la transmisión del BCH del terminal. El estado de los slots puede ser OCUPADO o LIBRE: es OCUPADO si se ha recibido

correctamente un paquete o ha sido el propio terminal el que lo ha transmitido. En el caso de que el slot esté marcado como OCUPADO, el FI también contiene la identidad del terminal transmisor.

Basándose en los FIs recibidos, cada terminal marca un slot, digamos el slot k, bien como RESERVADO, si el slot k-*N* se encuentra OCUPADO en al menos uno de los FIs recibidos en los slots del k-*N* al k-1 o bien como DISPONIBLE en caso contrario. Un slot DISPONIBLE puede emplearse para intentar nuevos accesos. Tras acceder en un slot DISPONIBLE, el terminal j reconocerá tras *N* slots (una trama) la transmisión como correcta si el slot es marcado como "OCUPADO por el terminal j" en todos los FIs recibidos o como incorrecta en el resto de casos. En la Fig. 1 se muestra un ejemplo de los FIs transmitidos por un grupo de terminales. Denominamos a la unión de los clusters a un salto (one-hop clusters: OH) que tienen un subconjunto común, como clusters a dos saltos (two-hop clusters: TH). Los terminales que pertenecen al mismo OH-cluster ven el mismo estado (DISPONIBLE o RESERVADO) para todos los slots; los terminales que pertenecen a distintos OH-clusters del mismo TH-cluster marcan como RESERVADO todos los slots usados en el TH-cluster mientras que terminales que pertenezcan a distintos OH-clusters normalmente ven un estado diferente. Como resultado de esto, los slots pueden reutilizarse en OH-clusters disjuntos, pero no en el mismo TH-cluster, y, por lo tanto, el problema del terminal oculto no puede producirse [4].

El BCH proporciona un canal broadcast a un salto fiable que puede ser usado tanto para señalización como para datos de usuario. A partir de esta base, pueden establecerse de una manera efectiva comunicaciones punto a punto entre los distintos terminales aprovechando la señalización distribuida que proporcionan los FIs. Con esta finalidad, cada entrada del FI incluye un flag de punto a punto (PointToPoint: PTP) que se emplea del siguiente modo:

- Un terminal pone el flag PTP de un slot dado en el FI si el paquete recibido en el slot es broadcast o el destino era el propio terminal.

De este modo, para establecer comunicaciones punto a punto pueden usarse todos los slots DISPONIBLES y también los slots RESERVADOS que cumplan lo siguiente:

- El flag PTP está a 0 en todos los FIs recibidos.
- El FI recibido del terminal destino marca el slot como LIBRE.

Estas condiciones permiten que las transmisiones punto a punto compartan el mismo slot cuando no hay colisión en los receptores.

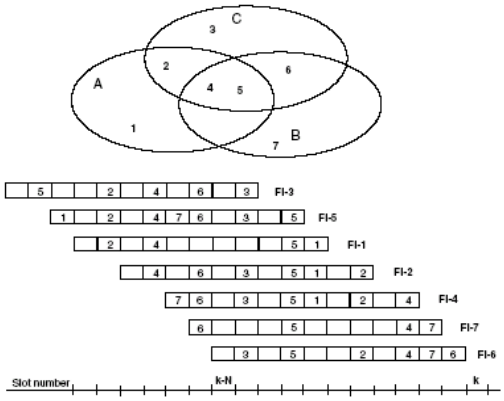


Figura 1: Ejemplo de los FI propagados por los terminales 1-7 en los cluster de un salto A, B y C representados por elipses.

Este hecho puede entenderse más claramente en los cuatro casos de la Fig. 2. Los casos a y b consideran dos terminales, 1 y 2, que pertenecen a distintos clusters no disjuntos. Asumiendo que el terminal 1 tiene ya activado un canal PTP con 3, el terminal 2 puede usar el mismo slot que el 1 aunque esté marcado como RESERVADO. De hecho, el único flag PTP activo es el transmitido por 3, y no recibido por 2 (cumpliendo la primera condición) y el FI generado por 4 marca el slot como LIBRE (cumpliendo la segunda condición). En el caso b, el FI generado por el terminal 3 y recibido por el 2 evita que el terminal 2 transmita (al no cumplirse la condición 1). En este caso, la transmisión paralela interferiría de hecho al terminal 3. En los casos c y d los dos terminales pertenecen al mismo cluster. En el caso d, el terminal 3 puede usar un slot reservado ya que se cumplen las dos condiciones (de hecho, es el caso del terminal expuesto) mientras que en el caso c no se cumple la segunda condición y habría colisión en el terminal 4. De todos modos, si se producen varios accesos simultáneos, puede seguir habiendo colisiones. Para saber si la transmisión ha sido correcta, el terminal transmisor comprueba que el slot esté marcado OCUPADO en el FI del terminal destino.

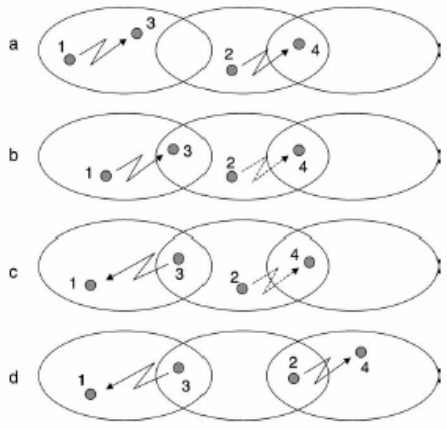


Figura 2: Ejemplo de transmisiones paralelas. Las transmisiones del terminal 1 se establecen primero. Las transmisiones permitidas para el terminal 2 se indican con línea continua.

2.2 Estrategias de asignación de ancho de banda

Una vez que un terminal ha adquirido un canal BCH, puede establecer canales de datos broadcast adicionales si el campo de datos del BCH no es suficiente. Del mismo modo, puede establecer canales PTP con sus distintos vecinos. En este artículo, sólo consideraremos comunicaciones extra PTP, así que de aquí en adelante, los canales adicionales son siempre referidos como PTP. Sin embargo, la estrategia propuesta puede ser generalizada, puesto que el dimensionado se realiza de acuerdo a las demandas de BCH.

En el modo de operación básico, cada slot de la trama puede usarse tanto para transmisiones PTP como BCH. Es esta situación, cuando crece el número de comunicaciones PTP, el número de slots DISPONIBLES para nuevos terminales que intenten acceder al sistema decrece, reduciendo el número de terminales que pueden acceder al sistema para un número dado de slots. Puesto que la adquisición de un canal básico de broadcast es obligatoria para acceder al sistema, un dimensionado apropiado de la red debe garantizar ciertos recursos para las transmisiones BCH. Como métrica para medir las prestaciones del BCH, consideramos la probabilidad de bloqueo. Un terminal se bloquea si no adquiere un canal BCH en un cierto número de tramas tras su aparición. Según esta situación, debe garantizarse un compromiso entre asegurar una probabilidad de bloqueo aceptable para los canales BCH mientras se proporciona el mayor throughput posible para las comunicaciones PTP. Para garantizar una probabilidad de bloqueo para nuevos terminales que acceden al sistema, proponemos una división de la trama en dos subtramas, donde las prestaciones del BCH no están limitadas por la cantidad de tráfico PTP en la red: una trama con N slots se divide en N_{BCH} y N_{PTP} slots para BCH y PTP respectivamente.

$$N = N_{BCH} + N_{PTP} \tag{1}$$

Para esta suposición, se requiere la sincronización temporal a nivel de trama y slot de cada terminal de la red, que puede obtenerse mediante GPS (Global Position System) u otras soluciones [7], [8]. Con esta subdivisión se consigue aumentar la probabilidad de acceder al sistema. Cuando un terminal intenta acceder al sistema, busca un slot DISPONIBLE. La existencia de un slot DISPONIBLE para un nuevo terminal sólo puede garantizarse estadísticamente: si los terminales vecinos tienen suficientes slots LIBRES, es probable que exista un slot común LIBRE para todos ellos. La subdivisión de la trama concentra los slots BCH LIBRES en la misma región haciendo más probable para un terminal nuevo encontrar un slot DISPONIBLE.

Si se emplea una subdivisión estática, N_{BCH} limita la densidad máxima de terminales que soporta el sistema. En el caso de que la densidad de terminales en la red sea menor, se desperdician recursos puesto que podrían establecerse canales PTP en los slots BCH libres. Por otro lado, si la densidad de terminales crece por encima de lo esperado, terminales que son bloqueados podrían acceder al sistema usando slots de la subtrama de PTP. Para superar estas limitaciones, se ha propuesto y evaluado una estrategia que desplaza el límite entre los slots dedicados a cada tipo de tráfico dentro de la trama.

En primer lugar, se define un conjunto de W posibles valores para N_{BCH} $\{N_1 < N_2 < \ldots < N_W\}$. El terminal i elige el valor $N_{BCH,i}$ dentro de este conjunto de acuerdo con la densidad de terminales ρ_i que observa. Hemos considerado dos posibilidades para medir esa densidad:

$$\rho_i = \left|NB_i\right| \tag{2}$$

donde NB_i es el conjunto de vecinos del terminal i, y su dimensión $|NB_i|$ es igual al número de canales BCH recibidos por este terminal.

$$\rho_i = \frac{1}{\left|NB_i\right|+1}\left(\sum_{j\in NB_i}\left|NB_j\right| + \left|NB_i\right|\right) \tag{3}$$

La ecuación (3) representa el número medio de vecinos en las cercanías del terminal i. Este valor puede obtenerse mediante la información transmitida en el FI por cada vecino del terminal i y los propios vecinos observados por dicho terminal. Según esta densidad, cada terminal actualiza el valor de $N_{BCH,i}$ cada trama y lo incluye en el FI que transmite a todos sus vecinos.

$$N_{BCH,i} = \begin{cases} N_1 & \text{if } \rho_i < th_1 \\ N_2 & \text{if } th_1 \le \rho_i < th_2 \\ \vdots & \\ N_{W-1} & \text{if } th_{W-2} \le \rho_i < th_{W-1} \\ N_W & \text{if } \rho_i \ge th_{W-1} \end{cases} \tag{4}$$

donde th_j representa la densidad máxima de terminales tolerada para un número de slots N_j de la subtrama BCH.

Del mismo modo que el terminal i envía este valor, $N_{BCH,i}$, recibe los correspondientes N_{BCH} de todos sus vecinos. Puesto que debe garantizarse que no se establezcan canales PTP en ninguna de las subtramas BCH de los vecinos, el número de slots donde el terminal i puede establecer canales PTP como transmisor, viene dado por:

$$N_{PTP-TX,i} = N - \max_{j\in NB_i}\left(N_{BCH,j}\right) \tag{5}$$

mientras que la subtrama donde puede recibir canales PTP solo está limitada por su propio $N_{BCH,i}$

$$N_{PTP-RX,i} = N - N_{BCH,i} \tag{6}$$

La Fig. 3 muestra un ejemplo de cómo pueden establecerse transmisiones PTP según el algoritmo de gestión de recursos. En la situación de la figura, el terminal 2 no puede establecer un canal PTP como transmisor con el terminal 1 en esa posición, aunque el slot esté DISPONIBLE, porque dicho slot pertenece a la subtrama BCH del terminal 3. Sin embargo, el terminal 1 puede transmitir al terminal 2 en ese mismo slot, puesto que esta transmisión no afecta al terminal 3. Este slot pertenece al conjunto de $N_{PTP\text{-}RX,2}$, pero no al de $N_{PTP\text{-}TX,2}$.

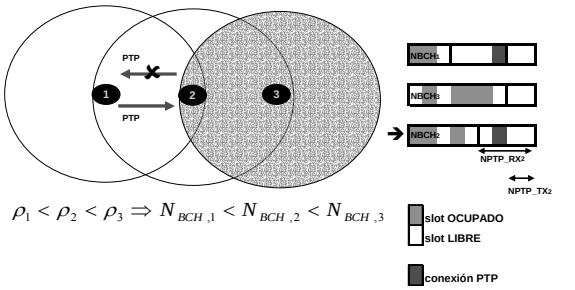


Figura 3: Ejemplo de cómo se pueden establecer transmisiones PTP según la división de la trama.

3 Evaluación de prestaciones

Se ha desarrollado un simulador que implementa todas las funcionalidades del protocolo de acceso al medio ADHOC MAC para evaluar las prestaciones de las estrategias de gestión de recursos y las prestaciones del servicio punto a punto sobre dicho protocolo. Puesto que el principal objetivo se centra en la evaluación de prestaciones, como primer nivel de análisis, se ha simplificado la capa física asumiendo que no hay desvanecimientos en el cálculo de la potencia recibida. La conectividad entre los terminales está determinada por la distancia entre los respectivos terminales. Como consecuencia, una transmisión, bien sea broadcast o punto a punto, sólo puede fallar por colisiones.

3.1 Evaluación de la asignación de ancho de banda

Las estrategias de asignación de recursos se han evaluado en una situación dinámica, donde los terminales se generan dentro de la red según un proceso de Poisson con tasa de llegada Y [nuevos terminales/s]. Cada terminal activo tiene un tiempo de vida exponencialmente distribuido con media L = 500 tramas, de modo que los parámetros Y y L definen el tráfico ofrecido por el servicio básico de broadcast. Los terminales se posicionan aleatoriamente en una región cuadrada de lado 1 Km. Bajo estas condiciones, las comunicaciones PTP se generan según un proceso de Poisson con intensidad X [conexiones PTP/s]. La fuente de cada conexión punto a punto es elegida aleatoriamente entre los usuarios con un BCH activo, mientras que el destino es elegido aleatoriamente entre los vecinos de la fuente. La duración de cada conexión punto a punto está exponencialmente distribuida con media D [tramas]. Los parámetros X y D definen el tráfico punto a punto ofrecido. Definimos un marco común para las simulaciones estableciendo la duración de la trama F = 100ms, el número de slots en una trama N = 30, el radio de cobertura R = 100 m y la duración media de las conexiones punto a punto D = 50 tramas. La modificación de estos parámetros de simulación tiene influencia sobre los valores absolutos de los resultados, mientras que los resultados comparativos presentados siguen siendo válidos.

En el modo de operación básico de ADHOC MAC, cada slot puede asignarse tanto para BCH como para conexiones punto a punto. Para acceder al sistema, cada usuario debe adquirir un canal BCH. Una vez adquirido, pueden establecerse comunicaciones adicionales. El número de tramas durante las que un terminal intenta acceder a la red adquiriendo un BCH se ha fijado en 10. Si transcurridas esas 10 tramas el terminal no ha conseguido adquirir un canal BCH, se considera bloqueado y abandona el sistema. Bajo estas condiciones, la Fig. 4 muestra la probabilidad de bloqueo de acceso al sistema frente al tráfico BCH ofrecido (equivalente a la densidad de usuarios que intentan acceder al sistema), definido como la densidad de canales BCH por slot, variando la intensidad del tráfico punto a punto. La probabilidad de bloqueo aumenta al crecer el tráfico punto a punto, puesto que se reduce el número de slots disponibles para conseguir un BCH. La capacidad del sistema puede entenderse como una capacidad BCH, definida como la capacidad de aceptar nuevos usuarios en el sistema, y la capacidad PTP, como el ancho de banda adicional para conexiones de datos. Los resultados de la Fig. 4 y la Fig. 5 muestran que es necesaria una estrategia de gestión de recursos más eficiente para repartir ambas capacidades. La división estática de la trama entre N_{BCH} y N_{PTP} slots intenta garantizar al menos el acceso al sistema de nuevos terminales. La Fig. 5 muestra la probabilidad de bloqueo de acceso al sistema con esta división estática para varios valores de N_{BCH} (10, 15, 20, 25 y 30). Se ha representado frente al número medio de vecinos en la red, que está directamente relacionado con la intensidad del tráfico BCH ofrecido.

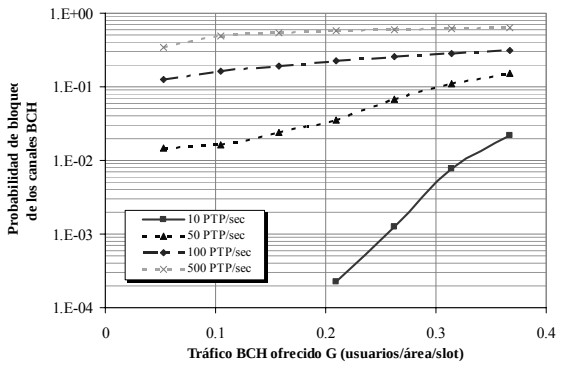


Figura 4: Probabilidad de bloqueo de acceso frente al tráfico BCH ofrecido cuando se varía la intensidad del tráfico PTP. Modo de operación básico.

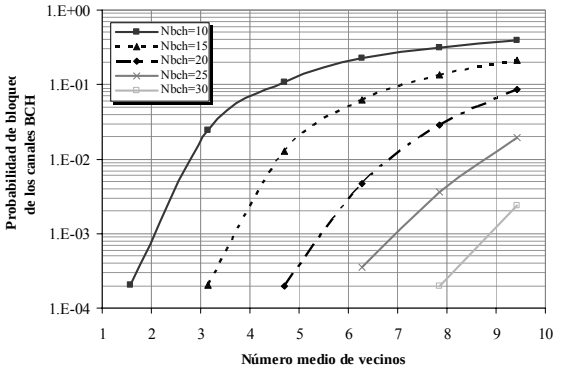


Figura 5: Probabilidad de bloqueo de acceso frente al número medio de vecinos con división estática de la trama.

Si se establece un determinado valor, por ejemplo un 1%, como un límite aceptable para la probabilidad de bloqueo, la Fig. 5 muestra que el N_{BCH} mínimo que garantiza este requerimiento cambia según el número de terminales en la red. Puesto que en una situación real esta densidad de usuarios en la red no va a ser conocida, el empleo de una subdivisión dinámica de la trama intenta optimizar el dimensionado de la red, realizándolo de manera local, según la densidad geográfica de terminales.

Los resultados de la Fig. 5 se han empleado como referencia para realizar el dimensionado estableciendo los valores th_j con $1 \leq j < W$ que determinan N_{BCH}. El conjunto de valores de N_{BCH} elegidos para esta subdivisión dinámica es $\{N_1=10, N_2=15, N_3=20, N_4=25, N_5=30\}$ y los valores elegidos para los umbrales *th* son $\{th_1=3, th_2=5, th_3=7, th_4=9\}$. Las decisiones se toman según (4). La Fig. 6 y la Fig. 7 muestran las prestaciones del algoritmo adaptativo. El uso del número medio de vecinos según (3) mejora claramente las prestaciones frente al uso del número propio según (2). Con el número medio, la probabilidad de bloqueo es menor y más estable para diferentes densidades de usuarios según la Fig. 6. Esto es confirmado en la Fig. 7, donde se observa que el número de slots asignado para BCH es mayor con (3). Además, el número de slots asignado para transmisiones PTP es también mayor, es decir, las diferencias entre $N_{PTP\text{-}TX}$ y $N_{PTP\text{-}RX}$ se reducen.

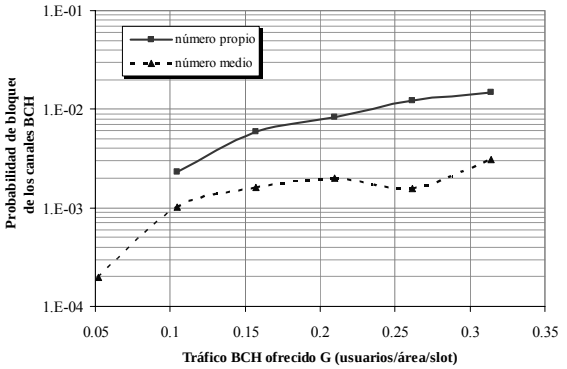


Figura 6: Probabilidad de bloqueo de acceso frente al número medio de vecinos con división dinámica de la trama usando el número propio y el número medio de vecinos.

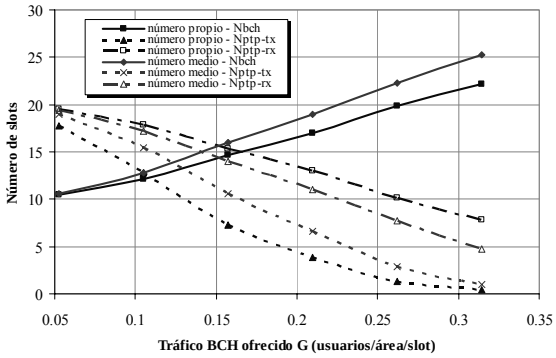


Figura 7: Número medio de slots asignados para BCH y PTP frente al número medio de vecinos con división dinámica de la trama usando el número propio y el número medio de vecinos.

El uso del número medio de vecinos permite a un terminal adaptar los límites de su trama a la variabilidad de la densidad de terminales en sus alrededores. Por ejemplo, la pérdida de un único vecino, que puede ser consecuencia de múltiples factores (movimiento, desvanecimientos, falta de batería, apagado temporal…) tiene un efecto menor sobre la densidad medida que si se emplea directamente el número propio de vecinos, consiguiendo que la subdivisión de la trama sea más estable.

Mediante la variación de los valores del conjunto de *th*, es posible ajustar el balance entre las comunicaciones BCH y PTP según los requerimientos de BCH. La Fig. 8 y la Fig. 9 muestran resultados similares para 3 conjuntos de umbrales diferentes {3, 5, 7, 9}, {3.5, 5.5, 7.5, 9.5} y {4, 6, 8, 10}. Para garantizar una probabilidad de bloqueo de acceso sobre el 1%, el conjunto {4, 6, 8, 10} puede ser suficiente, siendo la opción que proporciona un mayor ancho de banda para conexiones PTP.

3.2 Análisis de la eficiencia de las conexiones punto a punto

Una vez que se ha garantizado la probabilidad de bloqueo para nuevos accesos mediante la estrategia de asignación de recursos propuesta, la capacidad PTP restante debe ser gestionada de un modo eficiente. Como primer paso, se ha obtenido mediante simulación un valor máximo para esta capacidad. Estos resultados son válidos como la capacidad límite proporcionada por el protocolo, pero no podrán ser alcanzados en una situación dinámica, donde los requerimientos de BCH limitan la capacidad PTP real.

Para analizar la capacidad máxima del punto a punto sin interaccionar con el tráfico BCH, la simulación se ha realizado considerando una situación broadcast estacionaria, donde todos los terminales tienen un BCH activo. Con este propósito, al comienzo de la simulación se genera un número de terminales que estará activo durante toda la simulación. Este número define el tráfico ofrecido del servicio básico de

broadcast. Tras la generación, cada terminal intenta adquirir un BCH, de modo que tras un tiempo transitorio, todos los terminales han adquirido su BCH y ya se dispone del escenario estacionario. Por otro lado, también se han realizado simulaciones en una situación dinámica. Para estas simulaciones, se ha considerado el algoritmo adaptativo con el conjunto de umbrales {4, 6, 8, 10} y el número medio de vecinos, puesto que según la Fig. 8 y la Fig. 9 son los que proporcionan el mayor ancho de banda disponible para conexiones PTP garantizando un bloqueo aceptable para los nuevos accesos.

La Fig. 10 muestra, para ambos escenarios, el throughput máximo de las conexiones punto a punto, definido como la densidad máxima de transmisiones PTP correctas por slot, frente al tráfico broadcast ofrecido. En todos los casos, la cantidad de tráfico PTP ofrecido es suficiente para ocupar por completo los recursos disponibles. Se representan dos curvas para cada situación: la discontinua muestra las prestaciones de ADHOC MAC con el uso del flag PTP en los FIs, mientras que la continua se refiere al caso simplificado donde el flag PTP no se usa, es decir, sólo los slots DISPONIBLES se pueden emplear para accesos de tráfico punto a punto.

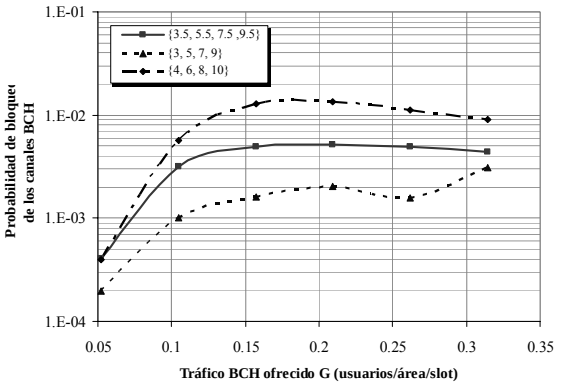


Figura 8: Probabilidad de bloqueo de acceso frente al número medio de vecinos con división dinámica de la trama (número medio de vecinos) para distintos umbrales.

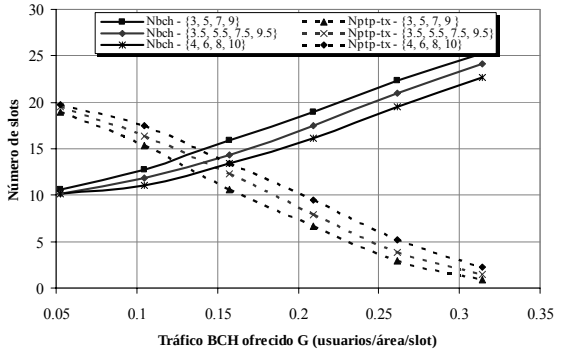


Figura 9: Número medio de slots asignados para BCH y PTP frente al número medio de vecinos con división dinámica de la trama (número medio de vecinos) para distintos umbrales.

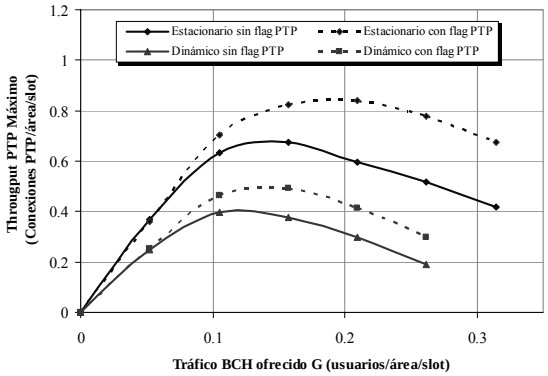


Figura 10: Throughput máximo para PTP frente al tráfico BCH ofrecido con y sin el flag PTP en los FIs. Tráfico BCH estático y dinámico.

ADHOC MAC proporciona un mayor reuso debido a las comunicaciones punto a punto paralelas en el mismo slot cuando se usa el flag PTP, puesto que éste soluciona el problema del terminal expuesto. A medida que aumenta el número de terminales (tráfico broadcast ofrecido) el uso del flag en el FI proporciona una creciente mejora respecto al caso en que no se usa. De hecho, un tráfico broadcast elevado implica una alta densidad de terminales y, en consecuencia, una alta probabilidad de situaciones de terminal expuesto. Además, en un escenario estacionario, el throughput máximo de punto a punto decrece si el tráfico ofrecido broadcast se incrementa por encima del valor 0.2 (terminales/área/slot), puesto que hay menos slots DISPONIBLES para punto a punto. Por debajo de ese valor, el throughput punto a punto crece con el tráfico BCH ofrecido, puesto que el número de conexiones punto a punto que pueden establecerse está limitado por el número de terminales dentro de la red. En una situación dinámica, puesto que hay slots que deben mantenerse LIBRES dentro de la subtrama BCH para garantizar el acceso de nuevos terminales, el throughput máximo que puede alcanzarse es menor que en un escenario estático sin división de trama, donde sólo los slots adquiridos para canales BCH no están DISPONIBLES para conexiones PTP.

4 Conclusiones

En este artículo, se han propuesto y evaluado mediante simulación estrategias de gestión de recursos para canales básicos de broadcast y conexiones de datos punto a punto en el protocolo ADHOC MAC. Ambos servicios pueden compartir de un modo eficiente los recursos totales mediante una subdivisión de la trama que permite realizar la gestión de manera independiente. Además, en una situación dinámica, la estrategia adaptativa propuesta, que realiza esta asignación según las densidades locales de terminales, proporciona un compromiso entre ambos servicios garantizando los requerimientos de acceso (probabilidad de bloqueo).

Respecto a las conexiones punto a punto, las prestaciones del protocolo son claramente mejoradas gracias al reuso de slots proporcionado por las transmisiones paralelas que pueden establecerse usando el flag PTP. A partir de estos resultados, en trabajos futuros se analizará la gestión de servicios punto a punto con distintos requerimientos de QoS.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia y Tecnología del Gobierno español y FEDER con el Proyecto TEC2004-04529/TCM

Referencias

[1] T. Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, New Jersey (1996).

[2] F. Borgonovo, A. Capone, M. Cesana, L. Fratta *ADHOC MAC: a new MAC Architecture for ad hoc Networks Providing Efficient and Reliable Point-to-Point and Broadcast Services*. Wireless Networks (WINET) Julio 2004, vol. 10, issue 4, pp. 359-366.

[3] CarTALK2000 Home page: www.cartalk2000.net

[4] M. Aoki, *Inter-vehicle communication: technical issues on vehicle control applications*. IEEE Communication Magazine. Octubre 1996, vol. 34, pp. 90-93.

[5] F. Borgonovo, L. Campelli, M. Cesana, L. Coletti, *MAC for ad-hoc inter-vehicle network: service and performance*. Proc. IEEE VTC fall 2003, Orlando, USA. Vol. 5, pp. 2789-2793

[6] F. Borgonovo, L. Campelli, M. Cesana, L. Fratta, *Impact of User Mobility on the Broadcast Service Efficiency in ADHOC MAC Protocol*. Proc. IEEE VTC 2005 Spring, Stockholm Sweden (2005).

[7] A. Ebner, H. Rohling, R. Halfmann, M. Lott, *Synchronization in ad hoc networks based on UTRA-TDD*. Proc. IEEE PIMRC 2002, Lisboa, Portugal.

[8] A. Ebner, H. Rohling, M. Lott, and R. Halfmann, *Decentralized slot synchronization in highly dynamic ad hoc networks*. Proc. WPMC 2002, Hawaii, USA.

Control de Admisión Óptimo en Redes Móviles Celulares con Predicción de Movimiento

José Manuel Giménez Guzmán, Jorge Martínez Bauset, Vicent Pla Boscà y Vicente Casares Giner
Departamento de Comunicaciones. Universidad Politécnica de Valencia
ETSI Telecomunicación. Camí de Vera s/n.
46022 - Valencia
Teléfono: 963 87 97 33, Fax: 963 87 73 09
E-mail: jogiguz@doctor.upv.es, (jmartinez,vpla,vcasares)@dcom.upv.es

Abstract *In this paper we study the impact of incorporating handover prediction information into the session admission control process in mobile cellular networks. The comparison is done between the performance of optimal policies obtained with and without the predictive information. A prediction agent classifies mobile users into two classes, those that will probably produce a handover to and/or from the cell under study and those that probably will not produce such handover. Moreover, the time instant has also been used in the prediction by means of a deterministic size window. We consider the classification error by modeling the false-positive and non-detection probabilities. Two different approaches to compute the optimal admission policy were studied: dynamic programming and reinforcement learning. Results show significant performance gains when the incoming predictive information is used in the admission process, and that higher gains are obtained when temporal information is used.*

1. Introducción

El Control de Admisión de Sesiones (CAS) es un aspecto clave en el diseño y funcionamiento de las redes móviles celulares multiservicio que ofrecen garantías para la calidad de servicio (QoS). Los parámetros de mérito que definen la QoS del sistema son de dos tipos: de nivel de paquete (como retardo, jitter o pérdidas) y de nivel de sesión (como las probabilidades de bloqueo de sesiones nuevas y de traspaso). Debido a la movilidad de los terminales, la disponibilidad de recursos al iniciar una sesión no garantiza que existan recursos suficientes para mantener la calidad de la misma mientras esté en curso.

Este artículo evalúa el impacto que, utilizar información de predicción del movimiento de los terminales en el proceso de CAS, tiene sobre los parámetros de QoS del nivel de sesión. Este estudio se hace desde una perspectiva de optimización, ya que consideramos que ésta no ha sido suficientemente explorada. Una de las ventajas del empleo de técnicas de optimización es que permite determinar el límite teórico para la ganancia que puede esperarse, no pudiendo éste establecerse mediante enfoques heurísticos.

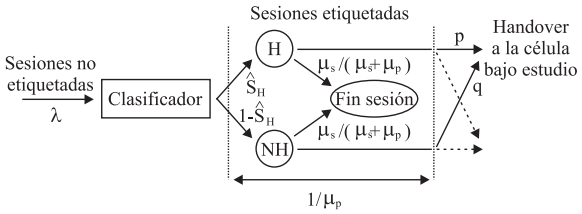
En los sistemas que no hacen uso de información predictiva en el CAS, se han utilizado tanto aproximaciones heurísticas como de optimización para mejorar las prestaciones del CAS. El tratamiento del CAS como un problema de optimización en entornos monoservicio ha sido estudiado en [1, 2] y en sistemas multiservicio en [3, 4]. Por otra parte, en sistemas que hacen uso de información predictiva, la mayoría utilizan enfoques heurísticos y se centran en un escenario monoservicio. Véase por ejemplo [5] y sus referencias.

Esta trabajo ha sido motivado, en parte, por el estudio presentado en [5]. De forma muy resumida, los autores proponen un sofisticado mecanismo de predicción de movimiento y un esquema que permite utilizar los resultados del mismo por el CAS, mejorando con ello las prestaciones del sistema. Una de las novedades que introduce el estudio es la de considerar, no sólo la predicción de traspasos entrantes a una célula, sino también los salientes. Ello se justificaría con el siguiente razonamiento. Considerar sólo los traspasos entrantes a una célula supondría reservar demasiados recursos, ya que durante el tiempo que trascurre desde que se predice el traspaso y se reservan recursos para el mismo, hasta que éste efectivamente ocurre, pueden haber ocurrido traspasos salientes y, por tanto, haber aumentado el número de recursos libres, haciendo la reserva anterior innecesaria.

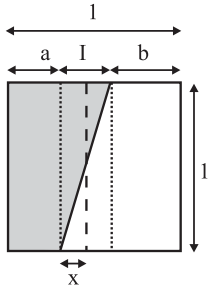
Este artículo utiliza parte del trabajo presentado en [6] para darle mayor coherencia, pero incorpora nuevas aportaciones. La primera es la evaluación comparativa del impacto que, sobre la QoS, tiene predecir la ocurrencia de traspasos sólo entrantes, sólo salientes, o entrantes y salientes de forma conjunta y facilitar dicha información al sistema de CAS. En [6] se evalúa sólo la ganancia que se obtiene al incorporar al CAS información relativa a la predicción de traspasos entrantes. La segunda aportación es la evaluación del impacto que tiene, sobre los parámetros del nivel de sesión, predecir no sólo la ocurrencia futura de traspasos sino los instantes futuros en los que éstos ocurrirán. El predic-

tor utilizado en [6] sólo predice la ocurrencia futura de traspasos entrantes pero no los instantes de ocurrencia. La tercera, es la aplicación de técnicas de optimización basadas en aprendizaje reforzado (*reinforcement learning*). Aunque estas técnicas ya se aplicaron en [6], aquí se justifica su validez al comparar sus resultados con los de una aproximación exacta como es la programación dinámica.

El resto del artículo está estructurado del siguiente modo. En la Sección 2 se describe el modelo del sistema y el del agente que predice la llegada de traspasos a la célula bajo estudio. Los dos algoritmos de optimización empleados se presentan en la Sección 3. La evaluación numérica, cuando se hace predicción sólo de entrada, se hace en la Sección 4. En la Sección 5 se introduce el predictor de salida y se evalúan las prestaciones cuando se utiliza sólo predicción de salida y de salida y entrada de forma conjunta. Se describe un modelo de predicción más complejo en la Sección 6, que predice los instantes en los que ocurrirán los traspasos. Finalmente, las conclusiones del artículo se presentan en la Sección 7.



(a) Diagrama de funcionamiento



(b) Incertidumbre en la predicción

Figura 1: Modelo del AP de entrada.

2. Descripción del modelo

Como modelo del sistema celular, se considera una única célula y sus inmediaciones, donde la célula tiene un total de C unidades de recurso, siendo su significado físico dependiente de la implementación tecnológica del interfaz radio. Se ofrecen un total de N servicios diferentes. Para cada servicio, hay dos flujos de llegadas: de nuevas sesiones y de traspasos de otra célula. Por tanto, hay N servicios y $2N$ tipos de llegadas diferentes.

Se considera más molesto para el usuario el bloqueo de un traspaso que el bloqueo de una nueva sesión, por lo que al realizar la asignación de recursos, los traspasos tiene prioridad sobre las sesiones nuevas. Por tratabilidad matemática, se asumen llegadas de Poisson y distribución exponencial para el tiempo de residencia y para la duración de las sesiones. La tasa de llegadas para sesiones nuevas (traspasos) del servicio i es λ_i^n (λ_i^h) y una sesión del servicio i consume b_i unidades de recurso, $b_i \in \mathbb{N}$. Para el servicio i, las tasas de duración de las sesiones y los tiempos de residencia son μ_i^s y μ_i^r respectivamente. El tiempo de ocupación de los recursos en una célula está también distribuido exponencialmente con tasa $\mu_i = \mu_i^s + \mu_i^r$.

2.1. Agente predictor de traspasos entrantes

Dado que el objetivo de nuestro estudio no es el diseño de un agente predictor (AP) sino el uso de la información que éste proporciona, se utilizará un modelo de éste, que se describe a continuación. El AP de entrada informa en todo momento al algoritmo de CAS acerca del número de terminales activos en las proximidades de la célula bajo estudio que se prevé van a realizar un traspaso hacia ésta.

En el momento en el que un terminal activo entra en las inmediaciones de la célula o uno que ya lo estaba pasa a estar activo, el AP lo clasifica como que probablemente va a producir un handover (H) o como que probablemente no lo hará (NH) atendiendo a las características del terminal (posición, velocidad, trayectoria, perfil histórico, ...) y otra información (mapa de calles y carreteras, hora del día, ...). Estas fuentes de información son cada vez más accesibles a los operadores celulares. Por ejemplo, en [7] se describen estándares relacionados con el posicionamiento en redes celulares. Tras un tiempo aleatorio, que en nuestro modelo sigue una distribución exponencial, el destino final se concreta y bien acaba produciéndose el traspaso o esta posibilidad se descarta definitivamente, por ejemplo porque la sesión termina o porque el terminal se mueve a otra célula.

El modelo del AP de entrada se caracteriza mediante tres parámetros: el tiempo medio desde que se realiza la predicción hasta que esta predicción se concreta μ_p^{-1}, la probabilidad p de provocar un traspaso si el terminal se ha etiquetado como H y la probabilidad de q de provocar un traspaso si el terminal se ha etiquetado como NH. Nótese que en general $q \neq 1 - p$. El funcionamiento básico del AP se muestra en la Fig. 1(a).

Los valores de p y q están relacionados a través del modelo de la Fig. 1(b), donde se muestra un cuadrado de superficie unidad que representa la población de los terminales que van a ser clasificados por el AP. El área sombreada representa la fracción de los terminales que finalmente realizarán un traspaso (S_H) y el área no sombreada el resto de terminales. El clasificador fija un umbral — representado mediante una línea discontínua vertical en la figura — que utiliza para discriminar entre aquellos terminales que probablemente

realizarán un traspaso — a la izquierda de la línea y cuya fracción denotados como $\hat{S}_H$— y los que no — a la derecha de la línea —. Como se ve en la figura, existe una zona de incertidumbre (I) que se representa mediante la línea oblicua de separación entre las zonas sombreada y no sombreada. Aunque por sencillez se ha utilizado un modelo lineal para la zona de incertidumbre, sería inmediato considerar un modelo diferente. El parámetro x representa la posición relativa del umbral del clasificador. Esta incertidumbre es la causante de los errores de clasificación: la zona triangular no sombreada a la izquierda del umbral (falso positivo, $\hat{S}^e_H$) y el triángulo sombreado a la derecha del umbral (no detección, $\hat{S}^e_{NH}$). A partir de la Fig. 1(b) se puede obtener:

$$1-p=\frac{\hat{S}^e_H}{\hat{S}_H}=\frac{x^2}{I(2S_H-I+2x)}$$
$$q=\frac{\hat{S}^e_{NH}}{1-\hat{S}_H}=\frac{(I-x)^2}{I(2-2S_H+I-2x)}$$

3. Optimización del CAS

Para encontrar la política que minimiza la media del valor esperado del coste por unidad de tiempo (*Average Expected Cost Rate*, AECR) se hace uso de la teoría de los *Procesos de Decisión de Markov* (MDPs) [8]. Se consideran políticas estacionarias y deterministas, que asocian acciones a estados, $\pi : S \rightarrow A$, de forma que la acción a tomar es función únicamente del estado actual.

Si llamamos $c(\boldsymbol{x}, a)$ el coste incurrido al ejecutar la acción a en el estado $\boldsymbol{x}$ y suponemos que al ejecutar una política π el sistema evoluciona a través de los estados $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_t$ en el intervalo $[0, t]$, entonces el coste total acumulado en el intervalo es

$$w^\pi(\boldsymbol{x}_0, t)=\sum_{m=0}^{t} c\big(x_m, \pi(x_m)\big)$$

Si el entorno es estocástico entonces $w^\pi(\boldsymbol{x}_0, t)$ es una variable aleatoria. Cuando el sistema empieza en el estado $\boldsymbol{x}$ y sigue la política π, el coste medio por unidad de tiempo esperado, denotado por $g^\pi(\boldsymbol{x})$, viene dado por:

$$g^\pi(\boldsymbol{x})=\lim_{t\rightarrow\infty}\frac{1}{t}E\left[w^\pi(\boldsymbol{x}, t)\right]$$

Para los sistemas bajo estudio, no es difícil ver que para cada política estacionaria determinista la cadena de Markov implícita tiene una matriz de transición de probabilidades de tipo *unichain* y por lo tanto, el coste medio por unidad de tiempo esperado no varía con el estado inicial [8]. El objetivo será pues encontrar la política π que minimice g^π, la cual denominaremos política óptima. En el caso que nos ocupa, la función de coste medio esperado a minimizar será una suma ponderada de las tasas de pérdidas de cada flujo de llegada: de sesiones nuevas y de traspasos de los diferentes servicios. Esto es,

$$g^\pi=\sum_{i=1}^{N}(\beta^n_i P^n_i \lambda^n_i+\beta^h_i P^h_i \lambda^h_i)$$

donde β^n_i (β^h_i) es el coste incurrido al bloquear peticiones de sesiones nuevas (de traspaso) y P^n_i (P^h_i) es la probabilidad de pérdidas de peticiones de sesiones nuevas (de traspaso), ambas del servicio i. En general, $\beta^n_i < \beta^h_i$ para tener en cuenta que el bloqueo de una petición de traspaso es menos deseable que el bloqueo de una sesión nueva.

Para el cálculo de la política óptima se han empleado dos aproximaciones diferentes. La primera es la técnica de *programación dinámica* (DP) [8] conocida como *mejoras sucesivas de la política (Policy Improvement)*. El segundo método empleado se basa en la teoría del *aprendizaje reforzado (Reinforcement Learning)* (RL) [9], concretamente, se ha empleado el algoritmo de aprendizaje reforzado de coste medio propuesto en [10].

La programación dinámica nos ofrece una solución exacta y nos permite evaluar los límites teóricos de la ganancia en prestaciones que es posible esperar al incorporar predicción de movimiento al proceso de CAS. Por otra parte, el aprendizaje reforzado nos permite manejar la explosión del espacio de estados y, además, ofrece la importante ventaja de ser un método que no necesita un conocimiento completo del modelo (*model free*).

3.1. Monoservicio

Representemos el estado del sistema por (i, j), donde i es el número de sesiones activas en la célula bajo estudio y j es el número de sesiones activas en las inmediaciones etiquetadas como H. El conjunto de posibles estados del sistema es

$$S:=\Big\{\boldsymbol{x}=(i,j): 0\le i\le C;\ 0\le j\le C_p\Big\}$$

donde C_p representa el número máximo de sesiones que pueden etiquetarse como H en un determinado instante. Se empleará un valor grande para C_p de modo que no tenga un impacto significativo en los resultados. Para cada estado $(i, j), i < C$, el conjunto de acciones posibles es $A := \{a : a = 0, 1\}$, siendo $a = 0$ la acción que rechaza la petición de una nueva sesión y $a = 1$ la acción en la que se acepta. En los estados (C, j) únicamente es posible la acción $a = 0$. Las peticiones de traspaso se aceptarán mientras hayan recursos disponibles, debido a su mayor prioridad.

En la Fig. 2 se muestran las transiciones desde y hacia el estado (i, j). En la figura se ha introducido el parámetro λ'_h que representa la tasa de llegada de traspasos no previsto, y cuyo valor puede expresarse como

$$\lambda'_h=(1-\hat{S}_H)\frac{\mu_p}{\mu_p+\mu_s}q\lambda$$

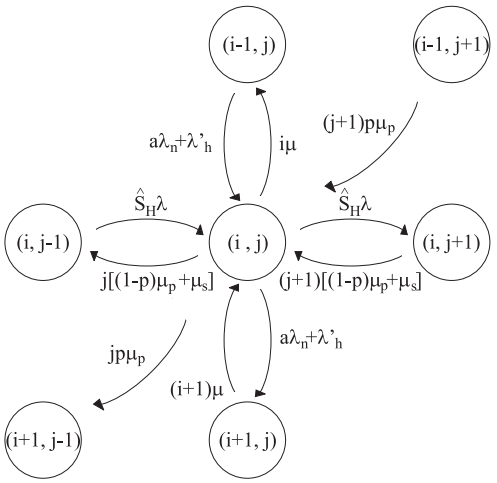


Figura 2: Diagrama de transiciones.

donde λ es la tasa de entrada al AP.

Las tasas de pérdidas pueden expresarse como

$$P_n\lambda_n = \sum_{\boldsymbol{x}:\ \pi(\boldsymbol{x})=0} \lambda_n p(\boldsymbol{x})$$

$$P_h\lambda_h = \sum_{\substack{\boldsymbol{x}=(C,j)\\ 0\le j\le C_p}} \big(\lambda'_h + jp\mu_p\big)p(\boldsymbol{x})$$

donde $p(\boldsymbol{x})$ es la probabilidad de estar en el estado $\boldsymbol{x}$. Por tanto, el objetivo del problema de optimización es encontrar la política π^* que minimiza g^π.

Aplicando la técnica *(policy improvement)* [8, Sección 8.6] y utilizando cualquier política inicial, por ejemplo *Complete Sharing*, se obtiene la política óptima en un número finito — y generalmente pequeño — de iteraciones.

3.2. Multiservicio

En esta sección, se formula el problema de la optimización como un *Proceso de Decisión Semi-Markoviano* (SMDP) bajo el criterio de optimización AECR, más concretamente como un problema de minimización de g^π. Se selecciona uno de los $2N$ tipos de llegada como el más prioritario, siendo sus peticiones siempre admitidas mientras hayan suficientes recursos disponibles, y por lo tanto, no se toman decisiones para ese flujo. Los instantes de decisión son los instantes en los cuales llega una nueva sesión o un nuevo traspaso, excepto si son del flujo más prioritario. Puesto que no hay que tomar decisiones en los momentos en los que una sesión abandona el sistema o finaliza, únicamente los eventos de llegada —salvo los correspondientes a llegadas del servicio más prioritario— son relevantes al proceso de optimización.

El espacio de estados se define como

$$S := \{\boldsymbol{x} = (x_0, x_1, k) : x_0, x_1, k \in \mathbb{N}\}$$

donde x_0 ($x_0 \le C$) es el número de unidades de recurso ocupadas en la célula bajo estudio, x_1 ($x_1 \le C_p$) es el número de unidades de recurso ocupadas por las sesiones etiquetadas como H por el AP y k ($1 \le k \le (2N-1)$), es el tipo de llegada. En cada instante de decisión, el sistema tiene que escoger una acción del conjunto de posibles acciones $A := \{0 = \text{rechazar}, 1 = \text{aceptar}\}$.

Los costes se definen del siguiente modo. En cualquier instante de decisión, el coste incurrido al aceptar una sesión es cero y el coste de rechazar una petición de sesión nueva (traspaso) del servicio i es β_i^n (β_i^h). Además, entre dos instantes de decisión se pueden acumular un coste adicional si se producen rechazos del flujo de sesiones más prioritario.

Las ecuaciones de Bellman para la optimización del coste medio por unidad de tiempo son

$$h^*(\boldsymbol{x},a) = \min_{a\in A_x}\Big\{c(\boldsymbol{x},a) - g^*\tau(\boldsymbol{x},a) + \\ + \sum_{\boldsymbol{x}\in S} p_{\boldsymbol{xy}}(a)\min_{a'\in A_{\boldsymbol{y}}} h^*(\boldsymbol{y},a')\Big\}$$

donde $h^*(\boldsymbol{x},a)$ es el valor medio relativo de tomar la acción óptima a en el estado $\boldsymbol{x}$ y entonces continuar tomando acciones de forma óptima, $c(\boldsymbol{x},a)$ y $\tau(\boldsymbol{x},a)$ son, respectivamente, el coste medio por unidad de tiempo y el tiempo medio de residencia al tomar la acción a en el estado $\boldsymbol{x}$, y $p_{\boldsymbol{xy}}(a)$ representa la probabilidad de moverse del estado $\boldsymbol{x}$ al estado $\boldsymbol{y}$ al tomar la acción $a = \pi(\boldsymbol{x})$. La política *greedy* π^*, que toma siempre las acciones que minimizan es del tipo $h^*(\boldsymbol{x},a)$ *gain-optimal* [10].

Si los parámetros del modelo pueden conocerse, entonces la solución a las ecuaciones de Bellman puede obtenerse mediante técnicas como la programación lineal o la programación dinámica. En entornos donde el número de estados puede ser muy alto y el conocimiento de los parámetros del modelo puede ser complejo, el problema puede ser intratable. La alternativa propuesta es usar la técnica de aprendizaje reforzado, concretamente el algoritmo SMART (*Semi-Markov Average Reward Technique*) [10].

SMART estima $h^*(\boldsymbol{x},a)$ mediante simulación, usando un método de diferencia temporal (TD(0)). Si en el instante de decisión $(m-1)$ el sistema está en el estado $\boldsymbol{x}$, se toma la acción a y el sistema se encuentra en el estado $\boldsymbol{y}$ en el instante de decisión m, entonces actualizamos el valor relativo de la pareja estado-acción del siguiente modo:

$$h_m(\boldsymbol{x},a) = (1-\alpha_m)h_{m-1}(\boldsymbol{x},a) + \\ +\alpha_m\Big\{c_m(\boldsymbol{x},a,\boldsymbol{y}) - g_m\tau_m(\boldsymbol{x},a,\boldsymbol{y}) + \\ + \min_{a'\in A_{\boldsymbol{y}}} h_{m-1}(\boldsymbol{y},a')\Big\}$$

donde $c_m(\boldsymbol{x},a,\boldsymbol{y})$ es el coste acumulado entre dos instantes de decisión consecutivos, $\tau_m(\boldsymbol{x},a,\boldsymbol{y})$ es el tiempo transcurrido entre los dos instantes de decisión, α_m es la tasa de aprendizaje en el instante de decisión m

y g_m es el coste medio por unidad de tiempo estimado del siguiente modo:

$$g_m = \frac{\sum_{k=1}^{m} c_k\left(\boldsymbol{x}_{(k)}, a_{(k)}, \boldsymbol{y}_{(k)}\right)}{\sum_{k=1}^{m} \tau_k\left(\boldsymbol{x}_{(k)}, a_{(k)}, \boldsymbol{y}_{(k)}\right)}$$

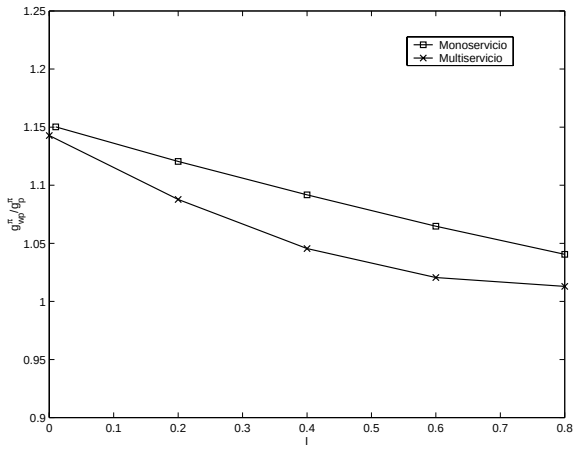
4. Evaluación numérica de la predicción de entrada

Para evaluar la mejora de prestaciones que supone incorporar la información predictiva en el proceso de CAS, se evalúa g_{wp}^{π}/g_{p}^{π}, donde el numerador (denominador) es el coste medio por unidad de tiempo esperado cuando no se usa (se usa) información predictiva. El escenario a estudiar consiste en la célula bajo estudio de radio r y las inmediaciones, siendo una corona circular de radio interior r y radio exterior $1.5r$.

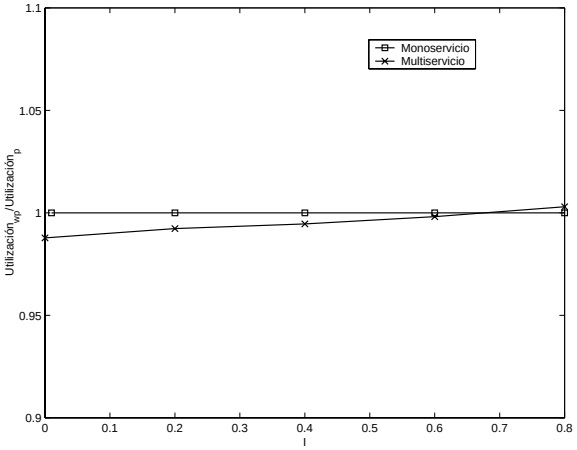
Los valores de los parámetros que definen el escenario son: $C = 10$ y $C_p = 60$ unidades de recurso, $N_h = \mu_i^r/\mu_i^s = 1$, $\mu_i^r/\mu_i^p = 0.5$, $S_H = 0.4$ y $x = U/2$. Para el escenario monoservicio ($N = 1$) se ha empleado $b_1 = 1$, $\lambda_1^n = 1$, $\mu_1 = \mu_1^s + \mu_1^r = 1$, $\beta_1^n = 1$ y $\beta_1^h = \beta = 20$. El valor de λ se escoge de modo que el sistema esté en equilibrio estadístico, es decir, la tasa con la que entran traspasos en una célula es igual a la tasa con la que salen. Se puede demostrar fácilmente que $\lambda = (1 - P_1^n)(1 - P_1^{tf})\lambda_1^n(N_h + \mu_1^r/\mu_1^p)(1/S_H)$. Nótese que en el escenario elegido para la evaluación numérica las tasas de llegada se eligen para obtener valores realistas de operación para la P_i^n y P_i^{tf} (P_1^n ($\approx 10^{-2}$) y P_1^{tf} ($\approx 10^{-3}$)). Para éstos valores, hacemos la aproximación $\lambda = 0.989\lambda_n(N_h + \mu_1^r/\mu_1^p)(1/S_H)$.

Para el escenario multiservicio empleamos $N = 2$ servicios, ocupando $b_1 = 1$ y $b_2 = 2$ unidades de recurso. La tasa de llegadas a la célula es $\lambda_1^{nc} = 0.8\lambda_T$, $\lambda_2^{nc} = 0.2\lambda_T$, donde $\lambda_T = 2$. La tasa de llegadas de sesiones nuevas a las inmediaciones (ng) y a la célula (nc) se relaciona mediante su respectiva relación de áreas, $\lambda_i^{ng} = 1.25\lambda_i^{nc}$. La tasa de llegadas de traspaso desde el exterior del sistema (ho) hasta las inmediaciones se calcula a partir de la relación de perímetros entre ese parámetro y la tasa de salida de traspasos desde la célula bajo estudio (hc), $\lambda_i^{ho} = 1.5\lambda_i^{hc}$. A partir del equilibrio de los flujos entrantes y salientes, podemos decir que $\lambda_i^{hc} = (1-P_i^n)(1-P_i^{tf})N_h\lambda_i^{nc}$, que a su vez puede aproximarse por $\lambda_i^{hc} = 0.989N_h\lambda_i^{nc}$. También se toman los siguientes valores $\mu_1 = \mu_1^s + \mu_1^r = 1$, $\mu_2 = \mu_2^s + \mu_2^r = 3$, $\beta_1^n = 1$, $\beta_2^n = 20$, $\beta_1^h = 10$ y $\beta_2^h = 200$. Para el algoritmo de aprendizaje reforzado se emplea una tasa de aprendizaje constante y de valor $\alpha_m = 0.01$ pero la tasa de exploración p_m se va reduciendo mediante $p_m = p_0/(1 + u)$, donde $u = m^2/(\gamma + m)$. Se ha empleado $\gamma = 1.0 \cdot 10^{11}$ con el fin de obtener $p_m = 1 \cdot 10^{-3}p_0$ cuando $m = 1 \cdot 10^7$. Se empieza con una tasa de exploración de $p_0 = 0.2$.

En todos los casos, se han comparado las prestaciones



(a) Ganancia de prestaciones al emplear predicción de traspasos.



(b) Ganancia en la utilización al emplear predicción de traspasos.

Figura 3: Comparación de prestaciones en escenario multiservicio al emplear predicción.

de la política óptima calculada al emplear predicción y al no hacerlo. En el caso en el que no se considera la predicción, la optimización se realiza ignorando la segunda componente del estado del sistema, es decir, el número de terminales clasificados como H.

En la Fig. 3(a) se muestra la variación de la ganancia g_{wp}^{π}/g_{p}^{π} conseguida al emplear predicción de traspasos entrantes con diferentes valores para la incertidumbre I. En el escenario multiservicio, para cada valor de I se han realizado 10 simulaciones con diferentes semillas, mostrándose la media. Como cabía esperar, el uso de la predicción conduce a una ganancia en las prestaciones en todos los casos, aunque dicha ganancia decrece conforme aumenta la incertidumbre (I) de la predicción.

Finalmente, hay que destacar que el reto principal en el diseño de técnicas de reserva de recursos para redes móviles celulares es el balance de dos requisitos conflictivos: reservar bastantes recursos con el fin de conseguir una baja probabilidad de terminación forzosa y mantener la utilización de los recursos alta bloqueando el menor número de peticiones de establecimiento de nuevas sesiones. La Fig. 3(b) muestra la variación de la ganancia en la utilización, esto es, el cociente entre

la utilización sin emplear predicción y la utilización al hacerlo, para diferentes valores de I. Estos resultados justifican la eficiencia de las políticas obtenidas, al reducir la probabilidad de terminación forzosa sin que ello suponga merma alguno en la utilización.

5. Predicción de salida

En esta sección se evalúa el impacto de la predicción de traspasos sólo salientes o entrantes y salientes de forma conjunta.

5.1. Agente predictor de traspasos salientes

La tarea de informar al algoritmo de CAS acerca del número de sesiones que se prevé van a realizar un traspaso hacia el exterior la llevará a cabo un nuevo elemento al que denominaremos AP de salida, cuyo modelo se presenta en la Fig. 4. De una manera similar a lo que ocurría con el AP de entrada, el AP de salida etiquetará las sesiones activas en función de si van a abandonar la célula (H^{out}) o no (NH^{out}), no estando esta clasificación exenta de errores, sino que, al igual que en el caso anterior, el error se modelará con las probabilidades de no detección y falso positivo.

La clasificación se hace cuando un terminal móvil activo entra en la célula bajo estudio o cuando un terminal, que ya está en la célula, pasa a estar activo. A diferencia del AP de entrada, el modelo del AP de salida se caracteriza únicamente mediante los parámetros p y q, cuyo significado es el mismo que en el AP de entrada, ya que el tiempo desde que se realiza la predicción hasta que se concreta, viene definido por el tiempo de residencia, que se considera distribuido exponencialmente con media $1/\mu_r$.

El funcionamiento básico del AP de salida se muestra en la Fig. 4. Los valores de p y q están relacionados mediante el mismo modelo que se tenía para el AP de entrada, es decir, el descrito en la Fig. 1(b). En el caso del AP de salida, el porcentaje de sesiones que efectivamente realizan un traspaso hacia el exterior de la célula (S_H^{out}) viene ahora determinado por la siguiente relación

$$S_H^{out} = \frac{\mu_r}{\mu_s + \mu_r} = \frac{N_H}{N_H + 1}$$

5.2. Evaluación numérica de la predicción de salida

Al igual que en el caso de la predicción de entrada, la evaluación de prestaciones se realizará comparando el coste medio por unidad de tiempo esperado sin información predictiva (g_{wp}^{π}) y al emplearla (g_p^{π}), del modo g_{wp}^{π}/g_p^{π}. Se ha aplicado al mismo escenario monoservicio descrito en la Sección 4 y con los mismos

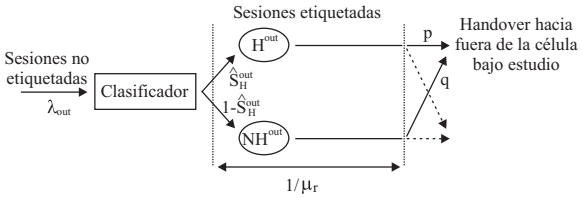


Figura 4: Diagrama de funcionamiento del AP de salida.

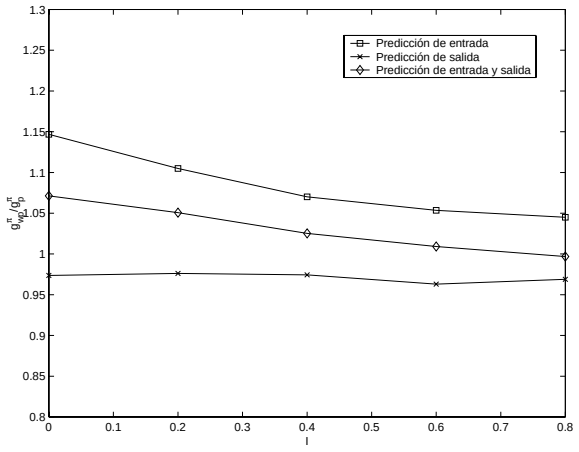


Figura 5: Ganancia de predicción al emplear predicción de salida, de entrada y de ambas.

parámetros, excepto que ahora se ha hecho con $\lambda = 2$ en vez de $\lambda = 1$ para acelerar la convergencia de la simulación, no siendo este cambio significativo en los resultados. La metodología de optimización empleada ha sido la del aprendizaje reforzado descrita en la Sección 3.2, con los mismos parámetros de configuración para el algoritmo SMART que se utilizaron en la Sección 4.

En el caso de emplear predicción de traspasos salientes, el espacio de estados estará definido por

$$S := \{\boldsymbol{x} = (x_0, x_{out}) : x_0, x_{out} \in \mathbb{N}\}$$

donde x_0 ($x_0 \leq C$) es el número de unidades de recurso ocupadas en la célula bajo estudio y x_{out} ($x_{out} \leq C$) es el número de unidades de recurso ocupadas por sesiones etiquetadas como H^{out} por el AP de salida.

También se han evaluado las prestaciones de un escenario en el que se disponga de ambos AP, el de entrada y el de salida. Para este caso, el espacio de estados vendrá dado por

$$S := \{\boldsymbol{x} = (x_0, x_{in}, x_{out}) : x_0, x_{in}, x_{out} \in \mathbb{N}\}$$

donde x_0 ($x_0 \leq C$) es el número de unidades de recurso ocupadas en la célula bajo estudio, x_{in} ($x_{in} \leq C_p$) es el número de unidades de recurso ocupadas por sesiones etiquetadas como H por el AP de entrada y x_{out} ($x_{out} \leq C$) es el número de unidades de recurso ocupadas por sesiones etiquetadas como H^{out} por el AP de salida.

En la Fig. 5 se muestra la variación de la ganancia g_{wp}^{π}/g_p^{π} conseguida con diferentes valores para la in-

certidumbre I, ejecutándose para cada valor de I 10 simulaciones con semillas diferentes y mostrándose la media. La gráfica representa las ganancias obtenidas al emplear predicción de salida, de entrada y salida, y únicamente de entrada. Como se observa, la ganacia cuando se utiliza predicción de salida es muy parecida a la de la Fig. 3(a). Por otra parte, y aunque no se muestra, también se ha comprobado que la utilización no se degrada significativamente al emplear este tipo de predicción. Como conclusión, se observa que el conocimiento del número de recursos que se van a liberar no es significativo para la optimización, siendo incluso independiente del valor de la incertidumbre I. En la gráfica se observa que incluso este conocimiento empeora ligeramente las soluciones, debido posiblemente a la mayor dificultad que supone optimizar en un espacio de estados más complejo que el existente al no emplear predicción. Se corrobora este hecho al analizar los resultados de la predicción de entrada y salida conjuntas, donde se ve que el incluir la información del AP de salida a la predicción de entrada empeora el proceso de optimización al aumentar el tamaño del espacio de estados.

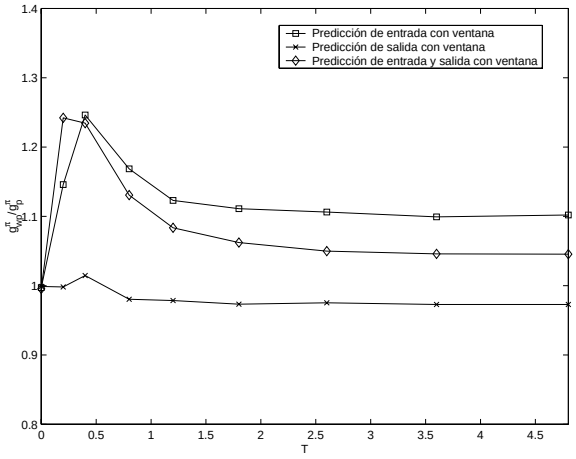


Figura 6: Ganancia de predicción al emplear predicción temporal.

6. Predicción temporal

Los modelos del AP introducidos hasta el momento no predicen los instantes en los que se van a producir los traspasos. En esta sección, el sistema de CAS conoce no el número de traspasos futuros sino cuántos se producirán en un tiempo inferior a T. Intuitivamente, parece obvio que tiene más relevancia para el sistemas de CAS el hecho de que el traspaso se vaya a producir de manera inminente que el hecho de que éste ocurra en un tiempo futuro indeterminado. Una aproximación similar se realiza en [5], donde la predicción de traspasos entrantes y/o salientes se realiza dentro de una ventana temporal de tamaño fijo.

En esta sección, los esquemas de funcionamiento de los AP descritos anteriormente son los mismos, excepto que el tiempo que trascurre desde que se realiza la clasificación hasta que el traspaso efectivamente ocurre es determinista y de valor T segundos.

6.1. Evaluación numérica de la predicción temporal

La evaluación numérica de este tipo de predicción, se ha hecho en el mismo escenario monoservicio empleado en la Sección 5.2. Los espacios de estados se definen del mismo modo a como se definían anteriormente, diferenciándose tres casos de predicción: de entrada, de salida, y de entrada y salida; todos ellos dentro de una ventana temporal. Así mismo, se ha empleado el algoritmo SMART, ejecutándose, para cada valor de T, 10 simulaciones con diferentes semillas y obteniéndose la media.

En la Fig. 6 se muestra el impacto sobre g^{π}_{wp}/g^{π}_{p} de la predicción, cuando ésta se realiza sobre un tiempo futuro T. En todos los casos, se ha tomado un valor para la incertidumbre del AP de entrada y/o salida de $I = 0.2$. Se aprecia como existe un óptimo para T, cuyo valor es cercano al tiempo medio entre llamadas (λ^{-1}), aunque posiblemente dependa también de otros parámetros.

Conforme T supera este valor óptimo, la ganancia se va reduciendo, posiblemente porque la información temporal se hace menos precisa. Cuando $T \to \infty$, la ganancia es, lógicamente, la misma que se obtiene cuando no se predicen los instantes en los que se van a producir los traspasos. Cuando T es inferior a su valor óptimo, la ganancia conseguida también disminuye, debido, posiblemente, al hecho de que el sistema no tiene tiempo de reaccionar. En el caso de que $T = 0$, no se produce ganancia, ya que no se hace predicción alguna.

Si se comparan las diferentes curvas entre sí, se observa como la información del AP de salida sigue sin aportar información significativa. Para valores de T próximos al óptimo se tienen ganancias similares cuando se utiliza predicción sólo de los traspasos de salida y predicción de los traspasos de entrada y salida conjuntamente. Además esta ganancia es sensiblemente superior a la que se produce cuando no se predice el instante de ocurrencia del traspaso. Finalmente, se ha comprobado que la utilización tampoco se ve afectada por el empleo de este esquema de predicción. Como conclusión, parece claro que cuando se utilizan técnicas de optimización para determinar la política óptima, la información del número de traspasos salientes no parece relevante.

Estos resultados parecerían contradecir las conclusiones de [5], pero allí la información de predicción se integra en el esquema de reserva de forma heurística, por lo que ésta aproximación y la de este trabajo no son, en principio, comparables.

7. Conclusiones

En este artículo se analiza la ganancia de prestaciones que puede obtenerse cuando se considera la información de predicción de traspasos al diseñar la política óptima de CAS en una red móvil celular multiservicio. Las prestaciones de una política se analizan mediante la suma ponderada de tasas de pérdidas de sesiones nuevas y de handover, ponderación que indica el mayor perjuicio que supone rechazar handovers respecto a sesiones nuevas. Las ganancias obtenidas se cuantifican como el cociente de esa suma ponderada al no emplear y al emplear información predictiva. Así, ganancias superiores a la unidad indicarán los beneficios de emplear información predictiva. El problema de la optimización se ha formulado como un proceso de decisión de Markov o de semi-Markov, para los cuales existen diferentes técnicas de resolución. En este caso se han empleado dos metodologías diferentes: programación dinámica y aprendizaje reforzado.

La información de predicción es aportada por un agente predictor, que etiqueta los terminales móviles activos en la célula, o en sus proximidades, que posiblemente ejecutarán un traspaso próximamente. Se han tratado diversos tipos de predicción por separado y conjuntamente. Por una parte, se ha empleado información de predicción del número de traspasos entrantes que van a tener lugar desde las inmediaciones de la célula bajo estudio hacia el interior de ésta, tanto en un entorno monoservicio como multiservicio. También se ha estudiado el impacto que sobre las prestaciones del sistema tiene incorporar predicción del número de terminales activos que van a salir de la célula bajo estudio. Por último, se ha evaluado también el impacto de disponer, además, de información más precisa del instante futuro en el que se producirá el traspaso.

Los resultados numéricos muestran que, con la aproximación de optimización utilizada, la información más útil es la de los traspasos que se prevé vayan a entrar en la célula bajo estudio. Además, precisar todavía más ésta información con el instante futuro en el que se producirá el traspaso, aporta substanciales ganancias adicionales. De este modo, se pueden conseguir ganancias de hasta un 25 % manteniendo la utilización del sistema con los esquemas de predicción estudiados, incluso con incertidumbres del 20 %.

Actualmente, se trabaja en el estudio e implementación de algoritmos de aprendizaje reforzado diferentes a SMART que aporten políticas de CAS más precisas y de una forma más rápida, incluso para espacios de estados más complejos. También se trabaja en la identificación de los parámetros del sistema que afectan al valor óptimo de T y en el análisis de sensibilidad de T frente a dichos parámetros.

Agradecimientos

El presente trabajo ha sido financiado por el *Ministerio de Educación y Ciencia* a través de los proyecto TIC2003-08272, TEC2004-06437-C05-01 y por el Programa de Incentivo a la Investigación de la Universidad Politécnica de Valencia.

Referencias

[1] R. Ramjee, R. Nagarajan, and D. Towsley, "On optimal call admission control in cellular networks," *Wireless Networks Journal (WINET)*, vol. 3, no. 1, pp. 29–41, 1997.

[2] N. Bartolini, "Handoff and optimal channel assignment in wireless networks," *Mobile Networks and Applications (MONET)*, vol. 6, no. 6, pp. 511–524, 2001.

[3] N. Bartolini and I. Chlamtac, "Call admission control in wireless multimedia networks," in *Proceedings of IEEE PIMRC*, 2002.

[4] V. Pla and V. Casares-Giner, "Optimal admission control policies in multiservice cellular networks," in *Proceedings of the International Network Optimization Conference (INOC)*, 2003, pp. 466–471.

[5] W.-S. Soh and H. S. Kim, "Dynamic bandwidth reservation in cellular networks using road topology based mobility prediction," in *Proceedings of IEEE INFOCOM*, 2004.

[6] V. Pla, J. M. Giménez-Guzmán, J. Martínez, and V. Casares-Giner, "Optimal admission control using handover prediction in mobile cellular networks," in *Proceedings of the 2nd International Working Conference on Performance Modelling and Evaluation of Heterogeneous Networks (HET-NETs '04)*, 2004.

[7] Y. Zhao, "Standardization of mobile phone positioning for 3g systems," *IEEE Communications Magazine*, pp. 108–116, July 2002.

[8] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 1994.

[9] R. Sutton and A. G. Barto, *Reinforcement Learning*. Cambridge, Massachusetts: The MIT press, 1998.

[10] T. K. Das, A. Gosavi, S. Mahadevan, and N. Marchalleck, "Solving semi-markov decision problems using average reward reinforcement learning," *Management Science*, vol. 45, no. 4, pp. 560–574, 1999.

Esquema Adaptativo de Reserva para Redes Móviles Celulares Multiservicio con Garantías de QoS

David García Roger, M.ª José Doménech Benlloch, Jorge Martínez Bauset y Vicent Pla
Departamento de Comunicaciones. Universidad Politécnica de Valencia
ETSI Telecomunicación. Camí de Vera s/n.
46022 - Valencia
Teléfono: 963 87 77 67, Fax: 963 87 73 09
E-mail: (dagarro,mdoben)@doctor.upv.es, (jmartinez,vpla)@dcom.upv.es

Abstract *We propose a novel adaptive reservation scheme designed to work in association with the well-known admission control policy Multiple Guard Channel (MGC). The scheme adjusts the MGC configuration parameters by continuously tracking the Quality of Service (QoS) perceived by users, adapting to any mix of aggregated traffic and handling both overload and underload situations. We provide two implementations of the scheme. The performance evaluation study confirms that our approach can satisfactorily deal with the nonstationarity nature of an operating network; besides the QoS objective is met with an excellent precision and the convergence period is much shorter than in previous proposals. These features along with its simplicity make our scheme superior to previous proposals.*

1. Introducción

El Control de Admisión de Sesiones (CAS) es un mecanismo clave en el diseño y operación de las redes móviles celulares multiservicio que ofrecen una calidad de servicio (QoS) garantizada. La movilidad de los terminales hace muy difícil garantizar que los recursos disponibles en el instante en que se establece la sesión, estarán también disponibles en las células visitadas mientras la sesión esté en curso.

Para el diseño del sistema de CAS se deben tener en cuenta no sólo parámetros asociados al nivel de paquete, como retardo máximo, *jitter* o probabilidad de pérdida, sino también parámetros del nivel de sesión, como probabilidades de bloqueo tanto de peticiones de nuevas sesiones como de peticiones de traspaso.

En este trabajo se estudia una novedosa estrategia de adaptación que funciona en cooperación con una política de CAS de la familia *trunk reservation* denominada *Multiple Guard Channel* (MGC) [1]. Esta política se caracteriza porque las decisiones de aceptar o rechazar una nueva sesión, por ejemplo del servicio r, dependen sólo del número de unidades de recurso libres en el sistema, donde el significado físico de una unidad de recurso depende de la tecnología específica que se haya seleccionado para implementar la interfaz radio.

En [1] las prestaciones de las políticas de CAS se evalúan determinando la tasa máxima agregada de peticiones de nuevas sesiones que puede ser ofrecida al sistema de forma que pueda garantizarse un determinado objetivo de QoS. A esta tasa agregada máxima se la denomina la capacidad del sistema. El objetivo de QoS se define en términos de cotas superiores para las probabilidades de bloqueo de peticiones de nuevas sesiones y de peticiones de traspaso.

Para la clase de políticas que son objeto de estudio, la capacidad del sistema es función de dos conjuntos de parámetros: los que describen los servicios como procesos de Markov y los que especifican el objetivo de QoS. Típicamente, durante la fase de planificación se considera que el primer conjunto de parámetros son estacionarios y, por tanto, un esquema de CAS estático (diseño en el caso peor). Sin embargo, es razonable prever que los esquemas estáticos de CAS no son adecuados en todas las situaciones, especialmente cuando el tráfico ofrecido varía con el tiempo. Para gestionar este escenario no estacionario, las aproximaciones comunes son las de estimar periódicamente los parámetros que describen los servicios o bien disponer de datos históricos del tráfico ofrecido.

Recientemente, se han propuesto numerosos esquemas adaptativos de CAS para redes móviles celulares. Dos ejemplos relevantes de esta aproximación en escenarios monoservicio son [2] y [3]. En [2] se ha propuesto un algoritmo de cuatro parámetros, basado en estimar la probabilidad de bloqueo percibida por los usuarios, para ajustar el número de *guard channels*. Para ello se define una ventana temporal de entre dos y diez horas durante la cual el sistema acumula información que le permite estimar la probabilidad de bloqueo de las peticiones de traspaso. Es obvio que esta ventana temporal es demasiado larga para capturar la dinámica de una red celular en operación. Además, el esquema propuesto en [2] no funciona correctamente para ciertos perfiles de tráfico cuando se utilizan los valores de los parámetros propuestos en [3], es decir, que el objetivo de QoS no se garantiza.

Un algoritmo de dos parámetros, basado en una adaptación probabilística, un concepto similar al del esquema *Random Early Detection* (RED), se propone en [3] para superar estas limitaciones. La ventaja principal de este esquema es que permite reducir la probabilidad de bloqueo de las peticiones de nuevas sesiones cuando el régimen permanente ha sido alcanzado y, por tanto, consigue aumentar la utilización de los recursos. No obstante, el periodo de convergencia es todavía del orden de las horas.

Esquemas de CAS adaptativos han sido estudiados también, por ejemplo, en [4], [5] y [6], tanto en escenarios monoservicio como multiservicio, pero en un contexto diferente al que se estudia en este trabajo. En éstos, el ajuste de la configuración de la política de CAS se basa en la tasa de llegada de las peticiones de traspaso. Ésta se estima en función del número de sesiones en curso en las células vecinas y del patrón de movilidad. En el contexto de este trabajo, el esquema de CAS ajusta la configuración de la política MGC usando una novedosa estrategia que se basa en monitorizar las decisiones de aceptación/rechazo tomadas y, en la propuesta que se describe, no incorpora ningún tipo de información de predicción. La configuración de una política de CAS especifica la acción (aceptación/rechazo) que debe ser tomada en cada estado del sistema cuando ocurre una petición de establecimiento de una nueva sesión o una petición de traspaso. El uso de información de predicción debería, obviamente, mejorar las prestaciones del algoritmo, pero su incorporación se deja para un estudio posterior. El esquema propuesto también utiliza una adaptación probabilística, como en [3], pero reduce considerablemente el periodo de convergencia.

El esquema adaptativo de CAS propuesto se diferencia de otras propuestas en: i) la simplicidad, ya que no requiere de intervalos de medida para estimar los parámetros de QoS; ii) la posibilidad de definir flujos "sensibles" (pudiendo establecer un orden de prioridades entre estos), y un flujo *best-effort* (utilizado para absorber la penalización que ineludiblemente ocurre durante episodios de sobrecarga); y iii) la elevada precisión con la que es capaz de cumplir el objetivo de QoS.

El resto del artículo se estructura de la siguiente forma: la Sección 2 describe el modelo del sistema, definiendo asimismo las políticas CAS usadas junto con el esquema adaptativo de reserva. La Sección 3 presenta las ideas básicas de funcionamiento del esquema adaptativo, detallando la estrategia de ajuste así como el soporte del multiservicio. Esto servirá de base para el desarrollo de la Sección 4, que describe la aplicación de estas ideas a un sistema celular multiservicio, mostrando los diferentes diagramas de funcionamiento que surgen en los diferentes niveles de acción. En la Sección 5 se presentan dos implementaciones del esquema adaptativo de reserva y los resultados obtenidos bajo condiciones de tráfico estacionario y no-estacionario. Finalmente la Sección 6 concluye el artículo.

2. Modelo de Sistema y Políticas de CAS Relevantes

Se estudia un escenario homogéneo en el que todas las células son estadísticamente idénticas e independientes. Por tanto, las prestaciones globales del sistema pueden ser analizadas concentrándose en una única célula. No obstante, el esquema adaptativo propuesto puede ser también utilizado en escenarios no homogéneos.

En cada célula un conjunto de R servicios diferentes compiten por C unidades de recurso. Para cada servicio se distinguen dos flujos de llegada, los de peticiones de nuevas sesiones y los de peticiones de traspaso, con lo que se definen $2R$ tipos de llegadas. Abusando un poco de la definición de un proceso de Poisson, diremos que para cada servicio r, $1 \leq r \leq R$, las peticiones de nuevas sesiones llegan de acuerdo a un proceso de Poisson cuya tasa es variable con el tiempo $\lambda_r^n(t)$ y solicitan c_r unidades de recurso por sesión. La duración de una sesión del servicio r (*unencumbered session duration*) está distribuida exponencialmente con tasa μ_r^s. El tiempo de residencia de una sesión del servicio r en una célula (*dwell time*) también está distribuido exponencialmente pero con tasa μ_r^d. Por tanto, el tiempo que una sesión del servicio r ocupa los recursos en una célula (*resource holding time*) está distribuido exponencialmente con tasa $\mu_r = \mu_r^s + \mu_r^d$.

Se considera que las peticiones de traspaso llegan según un proceso de Poisson cuya tasa varía con el tiempo $\lambda_r^h(t)$, y se supondrá que es una fracción conocida de la tasa de llegada de peticiones de nuevas sesiones $\lambda_r^n(t)$. Denotaremos por P_i, $1 \leq i \leq 2R$, las probabilidades de bloqueo de cada uno de los $2R$ flujos de llegada, siendo las probabilidades de bloqueo de las peticiones de nuevas sesiones $P_i^n = P_i$, mientras que las probabilidades de bloqueo de las peticiones de traspaso son $P_i^h = P_{R+i}$. El objetivo de QoS se expresa como cotas superiores para las probabilidades de bloqueo de peticiones de nuevas sesiones B_r^n y de peticiones de traspaso B_r^h. El estado del sistema se denotará por el vector $n \equiv (n_1, n_2, \ldots, n_{2R-1}, n_{2R})$, donde n_i es el número de sesiones en curso en la célula iniciadas como peticiones del flujo i. Denotaremos por $c(n) = \sum_{i=1}^{2R} n_i c_i$ al número de unidades de recurso ocupadas en el estado n.

Las políticas de CAS que son de interés en este trabajo son: i) *Complete-Sharing* (CS). Una petición es aceptada si hay suficientes unidades de recurso libres en el sistema. ii) *Multiple Guard Channel* (MGC). Sólo define un parámetro para el flujo de llegada i, $l_i \in \mathbb{N}$. Cuando una petición del flujo i ocurre en el estado n, se acepta si $c(n) + c_i \leq l_i$, o se bloquea en caso contrario.

La evaluación de las prestaciones del esquema adaptativo de CAS se ha realizado en cinco escenarios diferentes (A, B, C, D y E), que se definen en la Tabla 1, donde los parámetros de QoS B_i se han expre-

sado como valores porcentuales. Los parámetros de la Tabla 1 se han seleccionado para explorar posibles tendencias en los resultados numéricos, es decir, tomando el escenario A como referencia, el B representa el caso en el que la relación c_1/c_2 es menor, el C el que f_1/f_2 es menor, el D el que B_1/B_2 es menor y el E el que B_1 y B_2 son iguales. La tasa agregada de llegada de peticiones de nuevas sesiones se ha definido como $\lambda = \sum_{r=1}^{R} \lambda_r^n$, donde $\lambda_r^n = f_i\lambda$. La capacidad del sistema es la máxima λ (λ_{max}) que puede ofrecerse cumpliendo el objetivo de QoS.

Tabla 1: Definición de los escenarios estudiados

	A	B	C	D	E
c_1	1	1	1	1	1
c_2	2	4	2	2	2
f_1	0.8	0.8	0.2	0.8	0.8
f_2	0.2	0.2	0.8	0.2	0.2
B_1^n %	5	5	5	1	1
B_2^n %	1	1	1	2	1
	A,B,C,D,E				
B_r^h %	$0.1B_r^n$				
λ_r^n	$f_r\lambda$				
λ_r^h	$0.5\lambda_r^n$				
μ_1	1				
μ_2	3				

3. Fundamentos del Esquema Adaptativo de Reserva

Mientras que la aproximación al diseño de políticas de CAS mediante técnicas de optimización ha sido utilizada en escenarios estacionarios, en escenarios no estacionarios, en los que posiblemente se suceden los episodios de carga por encima (*overload*) y por debajo (*underload*) de la nominal, la QoS percibida por los usuarios puede ser considerablemente peor que la objetivo. En este trabajo se propone un esquema de reserva adaptativo que ha sido diseñado teniendo en cuenta dos aspectos claves para las prestaciones del mismo: el primero es conseguir un periodo de convergencia lo más corto posible, y el segundo es forzar cierto comportamiento del esquema durante los episodios de sobrecarga.

En la Fig. 1 se muestra la estructura general del esquema adaptativo de reserva propuesto. Como se puede ver, ante la llegada de una petición del flujo i, se decide la admisión o rechazo de la misma por parte del CAS; esta decisión se usará posteriormente para el ajuste de la configuración de la política de CAS por parte del esquema adaptativo.

3.1. Estrategia de Ajuste y Consideraciones sobre la Gestión de Escenarios Multiservicio

La mayoría de los esquemas adaptativos propuestos utilizan un estrategia de reserva basada en *guard channels* y se basan en aumentar el número de éstos para aquellos flujos de llegada que no cumplen el objetivo de QoS. Mientras que esquemas basados en este heurístico no son difíciles de intuir para escenarios monoservicio, para el caso multiservicio el problema es mucho más complejo. Una primera aproximación al problema podría consistir en clasificar los diferentes flujos de llegada en dos categorías genéricas: i) aquellos considerados como "sensibles" al tener definidos objetivos concretos de QoS; y ii) el *Best-Effort Stream* (BES), al no tener definidos objetivos concretos de QoS. Adicionalmente, el operador puede requerir la definición de prioridades entre los sensibles, de forma que se protejan con mayor efectividad aquellos que considere más importantes, como por ejemplo las peticiones de traspaso de los servicios que haya definido como *premium*. En este trabajo se supondrá que el operador puede definir, a su criterio, cualquier ordenación de los flujos en función de las prioridades deseadas. Para la política MGC, si $\mathbf{s} = (s_1, s_2, \ldots, s_{2R})$ es el conjunto de los flujos de llegada, la permutación $\pi^* \in \mathbf{\Pi}$, $\mathbf{\Pi} := (\pi_1, \ldots, \pi_i, \ldots, \pi_{2R}) : \pi_i \in \mathbb{N},\ 1 \leq \pi_i \leq 2R$, que es idéntica a la ordenación de los flujos deseada por el operador $\mathbf{s}^* = (s_{\pi_1}, s_{\pi_2}, \ldots, s_{\pi_{2R}})$ se denomina la "clasificación de prioridades", siendo s_{π_1} y $s_{\pi_{2R}}$ los flujos de llegada más prioritario (*Highest-Priority Stream*, HPS) y menos prioritario (*Lowest-Priority Stream*, LPS). Nótese que si existe un único BES, éste ocupará el lugar correspondiente al LPS dentro de la ordenación.

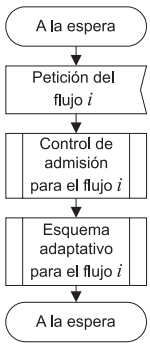
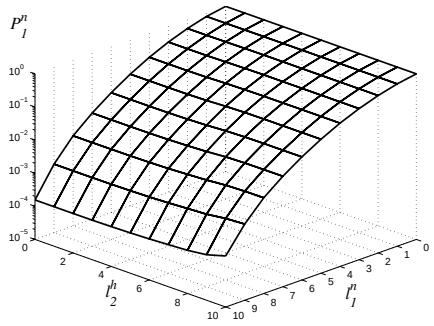


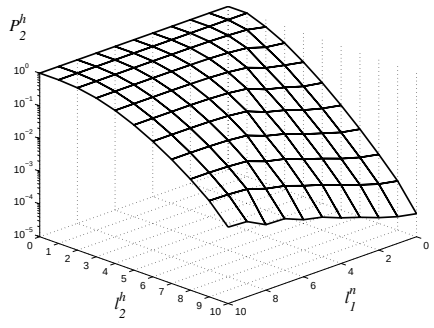
Figura 1: Funcionamiento del esquema adaptativo de reserva.

Se introduce a continuación nomenclatura adicional que se usará a lo largo del artículo. Se denota por s_r^n y s_r^h los flujos de peticiones de nuevas sesiones y de traspaso del servicio r, respectivamente y por s_i a un flujo de llegada genérico i. Respecto a los parámetros que definen la configuración de la política MGC, se denota por l_r^n y l_r^h los parámetros de configuración asociados con los flujos de llegada s_r^n y s_r^h, respectivamente y por l_i al parámetro de configuración asociado con el flujo de llegada genérico s_i.

El razonamiento heurístico que justificaría la estrategia de ajuste propuesta en este artículo es que ajustar el parámetro l_i, el parámetro de configuración asociado al flujo s_i, afecta directamente a la cantidad de recursos a la que tiene acceso el flujo i, e inversamente



(a) Flujo s_1^n



(b) Flujo s_2^h

Figura 2: Ejemplo de dependencia de la probabilidad de bloqueo de un flujo respecto a sus parámetros de configuración de la política CAS.

al resto de flujos. Así incrementar (disminuir) l_i, reduce (aumenta) P_i, y en contrapartida aumenta (reduce) las probabilidades de bloqueo del resto de flujos. A modo de ejemplo, la Fig. 2 muestra la dependencia de las probabilidades de bloqueo percibidas por s_1^n and s_2^h en el escenario A con $C = 10$ unidades de recurso, cuando se utiliza la política MGC y se ofrece una tasa agregada igual a la capacidad del sistema. La Fig. 2 (a) y (b) muestran el comportamiento de P_1^n y P_2^h cuando varían sus respectivos parámetros de configuración l_1^n and l_2^h, mientras los otros se mantienen constantes e iguales a sus valores óptimos. Como se observa, la validez del razonamiento heurístico enunciado anteriormente quedaría justificada, aunque con mayor precisión para P_1^n que para P_2^h.

3.2. Estrategia de Ajuste Probabilístico de los Parámetros de Configuración de la Política de CAS

Las estrategias de ajuste propuestas en [2, 3, 4, 5, 6] requieren una ventana temporal, en algunos casos denominada periodo de actualización (*update period*) durante la cual se acumula información para generar estimaciones de diferentes parámetros. La longitud del periodo de actualización suele ser constante [2, 3], pudiendo ser tan corta como sea posible [4], o estar asociada con la ocurrencia de un número específico de peticiones de nuevas sesiones [5].

Cuando se diseña un periodo de actualización constante, es necesario tener en cuenta ciertas consideraciones. Si es corto, el esquema de reserva se adaptará rápidamente a nuevas condiciones de tráfico, pero posiblemente consiguiendo unas prestaciones pobres. En cambio, si es largo conseguirá unas buenas prestaciones, pero posiblemente sea demasiado lento para gestionar la dinámica de una red en funcionamiento real. Además, debería tenerse en cuenta que las prestaciones del sistema dependerán de la precisión con la que se estimen sus parámetros. En [2] se sugiere que para conseguir una determinada precisión, cuanto más pequeñas sean las probabilidades de bloqueo, mayor deberá ser el periodo de actualización. Dado que las probabilidades de bloqueo objetivo son del orden de 10^{-2}–10^{-3}, el periodo de actualización acaba siendo excesivamente largo, del orden de horas [2].

El esquema de adaptación que se propone en este trabajo destaca por su simplicidad, dado que no requiere de intervalos de medida para estimar la QoS percibida por cada flujo. Suponiendo que los procesos de llegada de peticiones son estacionarios y que el sistema se encuentra en régimen permanente, parece intuitivo pensar que el esquema de ajuste no debería modificar los parámetros de configuración de aquellos flujos que cumplen los objetivos de QoS.

Basándonos en esta idea, proponemos realizar un ajuste probabilístico cada vez que ocurre una petición, es decir, cada vez que se toma una decisión de admisión/rechazo, de la siguiente forma. Supóngase que el objetivo de QoS del flujo de llegada i, B_i, se expresa como un número racional $B_i = b_i/o_i$, donde $b_i, o_i \in \mathbb{N}$. En régimen permanente es de esperar que si para el flujo i se verifica que $P_i = B_i$, éste experimenta, en media, b_i peticiones rechazadas y $o_i - b_i$ peticiones aceptadas, de un total de o_i peticiones ofrecidas. Por tanto, proponemos que cuando se tome la decisión de aceptar, se disminuya l_i con probabilidad $1/(o_i - b_i)$. En cambio, cuando se tome la decisión de rechazar, se aumente l_i con probabilidad $1/b_i$. Más concretamente, cuando se utilizan valores enteros para los parámetros de configuración l_i, como es el caso de la política MGC, el ajuste, cuando ocurre, consiste en sumar $+1$ ó -1 a l_i.

4. Funcionamiento del esquema adaptativo para CAS

En esta sección se describe la estructura de los bloques fundamentales del esquema adaptativo para el sistema de CAS presentado en la Fig. 1. En concreto, la Fig. 3 y la Fig. 4 muestran con detalle el bloque del CAS y el bloque del esquema adaptativo, respectivamente.

A efectos de este artículo, dos flujos de llegada, el HPS y el BES, requieren de un tratamiento específico. Por un lado, debido a su prioridad máxima, se tiene que poder cursar una solicitud del HPS siempre que haya suficientes unidades de recurso libres en el sistema

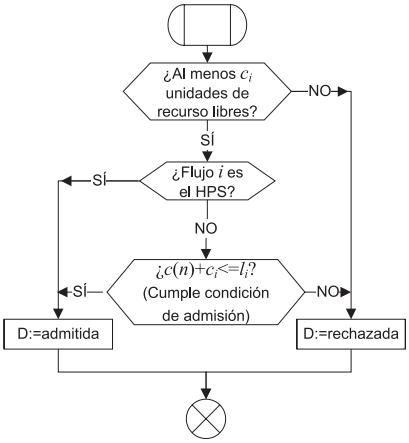


Figura 3: Descripción del bloque *Control de admisión para el flujo* i de la Fig. 2.

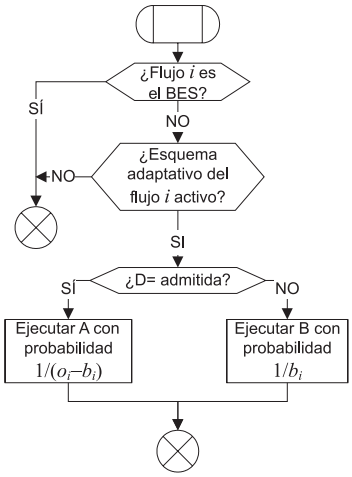


Figura 4: Descripción del bloque *Esquema adaptativo para el flujo* i de la Fig. 2.

(ver Fig. 3). Adicionalmente, no es necesario emplear un esquema adaptativo que ajuste la QoS percibida al objetivo en el caso del BES, puesto que éste carece de objetivos.

4.1. Bloque del CAS

El bloque del CAS para el parámetro de configuración asociado al flujo i se detalla en la Fig. 3. Inicialmente, la condición necesaria para admitir una petición del flujo i es la disponibilidad de suficientes unidades de recurso libres, (c_i). Nótese que, verificada la condición anterior, al HPS siempre se le acepta la petición. El resto de flujos, para ver aceptada su petición, necesitan cumplir la condición de aceptación asociada a la política MGC.

4.2. Bloque del esquema adaptativo

El bloque del esquema adaptativo para un flujo i se detalla en la Fig. 4 y es el encargado de que cada flujo de llegada reciba un trato diferenciado por parte del sistema. Un caso particular sucede cuando el LPS es en realidad un BES, que por no estar sujeto a objetivos de QoS, no requiere de un esquema adaptativo. El resto de flujos sensibles necesita tener un esquema adaptativo asociado al CAS para garantizar los objetivos de QoS. Por lo general dicho esquema adaptativo estará activo, aunque bajo ciertas circunstancias (que se explicarán en la Sección 4.3) podría ser necesario desactivarlo con el propósito de beneficiar a flujos de mayor prioridad.

4.3. Ajuste probabilístico del CAS efectuado por el esquema adaptativo

El ajuste probabilístico descrito en la Sección 3.2 necesita de mecanismos adicionales que posibiliten un correcto funcionamiento del sistema multiservicio en circunstancias en las que no se pueden mantener los objetivos de QoS para todos los flujos (por ejemplo en episodios de sobrecarga o cambios en los factores de penetración f_i). Es decir, se desea ofrecer una protección mayor a aquellos flujos con una prioridad superior dentro de la clasificación de prioridades. El "modo directo" de ofrecer protección es aumentar (probabilísticamente) l_i, de forma que cuando $l_i = C$, se proporciona acceso total a los recursos, pero hacer $l_i > C$ no proporciona ventaja alguna. En dichos casos se propone un "modo indirecto" de beneficiar a dicho flujo, consistente en limitar el acceso a recursos de uno (o varios) flujos de menor prioridad (mediante la disminución del parámetro de la configuración asociado l).

El trabajo descrito se puede considerar como una generalización de [9] al introducir la clasificación de prioridades, la existencia de un flujo *best-effort* y la penalización progresiva de los flujos de prioridad inferior al aumentar el tráfico ofrecido.

El ajuste del CAS frente al rechazo de una petición del flujo i se muestra en la Fig. 5. Nótese que es necesario desactivar los esquemas adaptativos de los flujos de prioridad inferior involucrados en el ajuste indirecto; si esto no se hiciera, los esquemas adaptativos eventualmente terminarían por anular de manera natural el ajuste efectuado. Nótese también que l_i puede ser mayor que C y que l_{π_k} puede ser negativo.

En la Fig. 6 se muestra cómo se ajusta el CAS frente a la admisión de una peticion del flujo i. Ahora los procedimientos realizados son los contrarios. Al contrario que antes, si ahora un flujo de prioridad inferior desactivado deja de tener prohibido el acceso a los recursos, se activa el esquema adaptativo del flujo de prioridad inmediatamente superior. Obsérvese asimismo que si el LPS es un BES, no existe esquema adaptativo que reactivar.

5. Evaluación de prestaciones

En esta sección se evalúan las prestaciones del esquema adaptativo de reserva asociado a la política

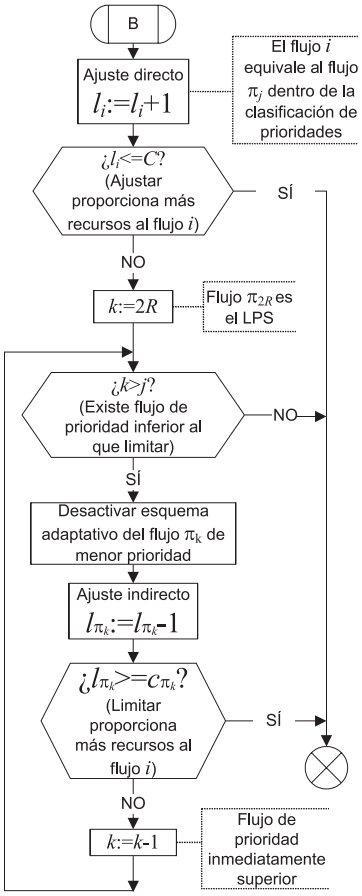


Figura 5: Algoritmo de ajuste de los parámetros del CAS tras una decisión de rechazo.

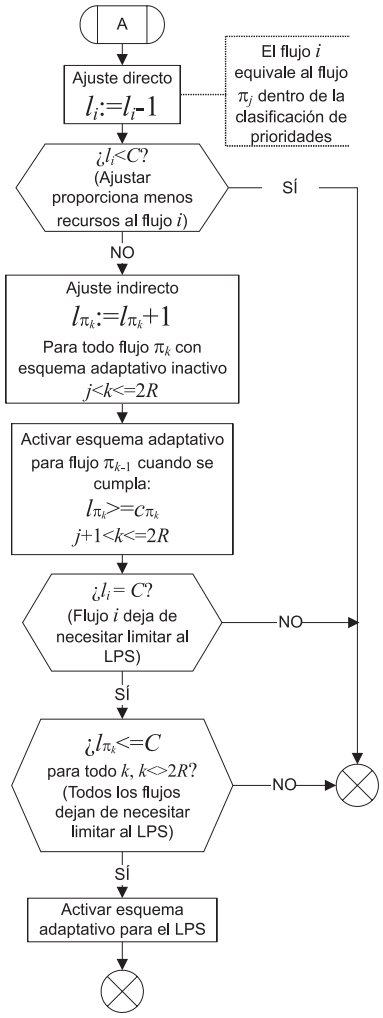


Figura 6: Algoritmo de ajuste de los parámetros del CAS tras una decisión de aceptación.

MGC. En el análisis de las prestaciones se hace uso de Möbius™ [7]. Möbius™ es una herramienta *software* que soporta *Stochastic Activity Networks* (SANs). Möbius™ permite simular las SAN que modelan el tipo de redes móviles celulares multiservicio con QoS de interés en este estudio, e incluso bajo ciertas condiciones resolver numéricamente las cadenas de Markov de tiempo continuo asociadas. En particular, el esquema propuesto cumple las condiciones para su resolución numérica.

La Tabla 2 muestra la capacidad del sistema en ausencia de política (CS) y cuando se emplea la política fija MGC (sin esquema adaptativo), para los cinco escenarios definidos en la Tabla 1, con $C = 10$ unidades de recurso y tráfico estacionario. Véase [1, 8] para más detalles acerca de cómo determinar la capacidad del sistema.

Tabla 2: Capacidad del sistema con $C = 10$ para diferentes políticas y escenarios.

	A	B	C	D	E
MGC	1.89	0.40	1.52	1.97	1.74
CS	1.54	0.37	1.37	1.74	1.54

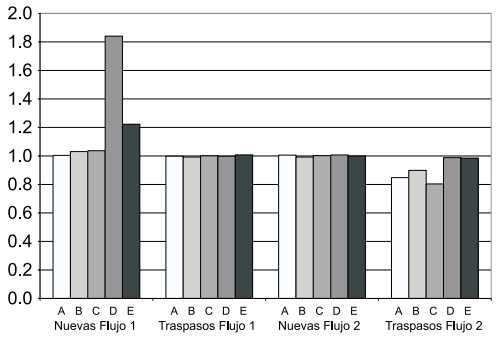
Supondremos para los escenarios de la Tabla 1 que la clasificación de prioridades fijada por el operador es $\mathbf{s}^* = (s_2^h, s_1^h, s_2^n, s_1^n)$.

Se analizarán mediante simulación dos implementaciones concretas del esquema adaptativo estudiado, una que considera al flujo s_1^n como *best-effort* y otra que lo considera como un flujo sensible, que de ahora en adelante se indicarán en el texto como "implementación con flujo *best-effort*" e "implementacion sin flujo *best-effort*".

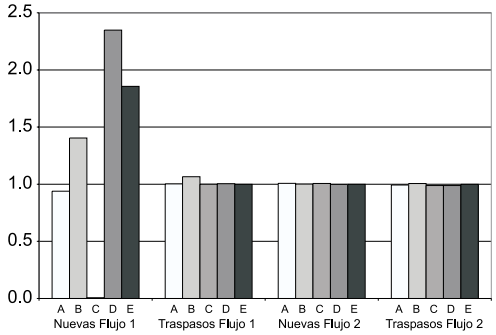
5.1. Prestaciones frente a tráfico estacionario

La Fig. 7 (a) y (b) muestra las prestaciones de las dos implementaciones del esquema adaptativo cuando se ofrece al sistema un tráfico estacionario e igual a la capacidad del sistema en cada escenario (ver Tabla 2).

Para proporcionar una perspectiva adicional la Fig. 8 muestra las probabilidades de bloqueo de los cuatro flujos con respecto al tráfico ofrecido para el escenario C con $C = 10$ unidades de recurso. La implementacion sin flujo *best-effort* (a), trata de ajustar la probabilidad de bloqueo percibida por los flujos a su ob-

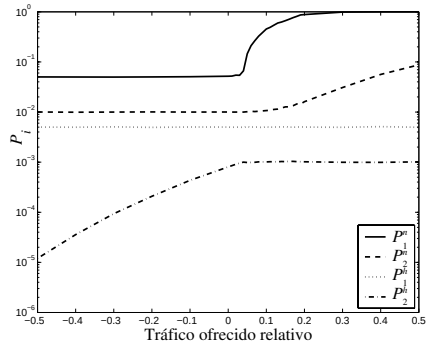


(a) Implementación sin flujo *best-effort*

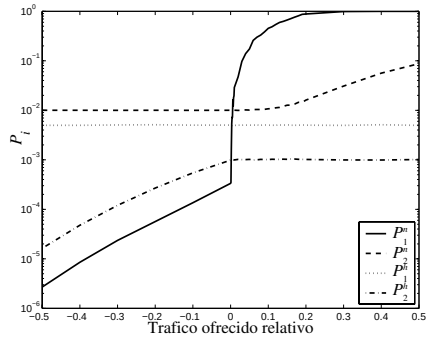


(b) Implementación con flujo *best-effort*

Figura 7: P_i/B_i para un sistema estacionario al que se ofrece λ_{max} (capacidad del sistema).



(a) Implementación sin flujo *best-effort*



(b) Implementación con flujo *best-effort*

Figura 8: P_i para un sistema estacionario y diferentes valores de λ (expresado como $(\lambda - \lambda_{max})/\lambda_{max}$).

jetivo. Como se puede apreciar, en el tramo de carga por debajo de la nominal ($\lambda/\lambda_{max} < 1$) el HPS (flujo s_2^h) se beneficia exclusivamente de la capacidad extra, mientras que en el tramo de carga por encima de la nominal ($\lambda/\lambda_{max} > 1$), se penaliza en primer lugar al LPS (flujo s_1^n) y (si es necesario) y en segundo lugar al flujo de prioridad inmediatamente superior (s_2^n) respetando así la clasificación de prioridades. Por otro lado, la implementación con flujo *best-effort* (b) se comporta de manera similar a la anterior en el tramo de carga por encima de la nominal, mientras que en el tramo de carga por debajo de la nominal los beneficiarios de la capacidad extra son tanto el HPS como el LPS.

En cualquier caso es interesante destacar que aunque durante los episodios de carga por debajo de la nominal el sistema tiende a rechazar más peticiones de las requeridas, existen flujos que se benefician de este excedente de capacidad.

5.2. Prestaciones frente a tráfico no estacionario

Se estudia el comportamiento transitorio de las probabilidades de bloqueo en el escenario A, al emplear el esquema adaptativo propuesto en su implementacion sin flujo *best-effort*, con la política MGC frente a tráfico no estacionario. Se fuerza un incremento repentino en el tráfico ofrecido, todo ello en un instante concreto del tiempo en el que el esquema adaptativo se considera en régimen permanente, soportando una $\lambda = 0.66\lambda_{max}$.

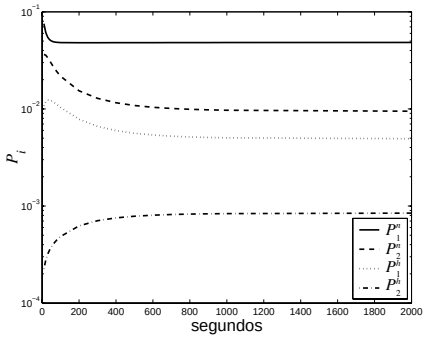


Figura 9: Comportamiento transitorio del esquema adaptativo usando la implementación sin flujo *best-effort*, con un incremento repentino del 50 %.

La Fig. 9 muestra el comportamiento transitorio de la probabilidad de bloqueo para un incremento repentino desde λ hasta λ_{max}. Como se observa, la velocidad de adaptación está en el rango de aproximadamente miles de segundos para el escenario estudiado. Esto representa una reducción substancial comparada con el tiempo de convergencia logrado por propuestas anteriores, que oscila entre 10 y 100 veces más [2, 3]. Nótese que cuando el tráfico ofrecido es superior al nominal, en parte se beneficia la convergencia del esquema adaptativo, porque se realizan más actualizaciones por unidad de tiempo. Por limitación de extensión

no se detallarán, pero existen mecanismos adicionales a incluir en este esquema adaptativo para establecer un compromiso adecuado entre la velocidad de convergencia del esquema adaptativo y la precisión con la que se desean lograr los objetivos de QoS.

6. Conclusiones

Se ha propuesto un esquema adaptativo de reserva que se puede usar en escenarios no estacionarios con una de las políticas de control de admisión de sesiones más empleadas en redes móviles celulares multiservicio, *Multiple Guard Channel*. Dos de los rasgos principales del esquema son su simplicidad y la habilidad para la monitorización y ajuste continuo de la QoS percibida por los usuarios.

Se proporcionan dos implementaciones diferentes del esquema, una en la que se considera que el flujo de menor prioridad tiene requisitos de QoS y otra en la que se le considera un flujo *best-effort*. En consecuencia en esta última implementación el flujo de menor prioridad disfruta una QoS impredecible.

Las simulaciones realizadas muestran que se logran los objetivos de QoS con una precisión excelente y que el periodo de convergencia es mucho más corto que en propuestas anteriores (aproximadamente del orden de miles de segundos), confirmando que este enfoque puede manejar de manera satisfactoria la no estacionariedad de una red en funcionamiento.

Aunque se deja como tarea futura un estudio comparativo de este esquema en situaciones de sobrecarga, se conjetura que se comportará substancialmente mejor que empleando otros esquemas adaptativos propuestos con anterioridad. Adicionalmente, otro trabajo futuro consistirá en el estudio del esquema adaptativo propuesto en escenarios no homogeneos. En los mismos, una posible mejora podría radicar en basar el ajuste no exclusivamente en las decisiones del CAS, sino también en información adicional como la predicción de los traspasos venideros. Asimismo se pretende extender este esquema adaptativo a otras familias de políticas de CAS, como aquellas para las que las probabilidades de estado estacionarias de la cadena de Markov de tiempo continuo tiene una forma producto.

Agradecimientos

Este trabajo ha sido financiado por el Gobierno Español y la Comisión Europea bajo los proyectos TIC2003-08272 y TEC2004-06437-C05-01 (30 % PGE y 70 % FEDER), por el Ministerio de Educación y Ciencia a través de la beca AP-2004-3332 y por la *Generalitat Valenciana* a través de la beca CTB/PRB/2002/267.

Referencias

[1] V. Pla, V. Casares-Giner, “Optimal admission control policies in multiservice cellular networks”, Proceedings of the International Network Optimization Conference (INOC), pp. 466-471, Paris, (France), 27-29 octubre 2003.

[2] Y. Zhang, D. Liu, “An adaptive algorithm for call admission control in wireless networks”, Proceedings of the IEEE Global Communications Conference (GLOBECOM), pp. 3628-3632, San Antonio, (USA), noviembre 2001.

[3] X.-P. Wang, J.-L. Zheng, W. Zeng, G.-D. Zhang, “A probability-based adaptive algorithm for call admission control in wireless network”, Proceedings of the International Conference on Computer Networks and Mobile Computing (ICCNMC), pp. 197-204, Shanghai, (China), 20-23 octubre 2003.

[4] O. Yu, V. Leung, “Adaptive Resource Allocation for prioritized call admission over an ATM-based Wireless PCN”, IEEE Journal on Selected Areas in Communications, pp. 1208-1224, Vol. 15, septiembre 1997.

[5] P. Ramanathan, K. M. Sivalingam, P. Agrawal, S. Kishore, “Dynamic Resource Allocation Schemes During Handoff for Mobile Multimedia Wireless Networks”, Journal on Selected Areas in Communications, pp. 1270-1283, Vol. 17, julio 1999.

[6] O. Yu, S. Khanvilkar, “Dynamic adaptive QoS provisioning over GPRS wireless mobile links”, Proceedings of the IEEE International Conference on Communications (ICC), pp. 1100-1104, Vol. 2, New York, (USA), 28 abril- 2 mayo 2002.

[7] Performability Engineering Research Group (PERFORM), Möbius^{TM}. User Manual. Version 1.6.0: http://www.perform.csl.uiuc.edu/mobius/manual/MobiusManual_160.pdf.

[8] D. García, J. Martínez, V. Pla, “Admission Control Policies in Cellular Multiservice Networks: Configuration, Performance and Sensitivity”, Lecture Notes in Computer Science, pp. 121-135, Vol. 3427, ISBN: 3-540-25329-7, ISSN: 0302-9743 (2005).

[9] D. Garcia-Roger, M.ª Jose Domenech-Benlloch, J. Martinez-Bauset, V. Pla, “Adaptive Admission Control Scheme for Multiservice Mobile Cellular Networks”, Proceedings of the 1st Conference on Next Generation Internet Networks (NGI2005), Roma, (Italy), 18-20 abril 2005.

Análisis de la estabilidad de modelos de movilidad en simulaciones de redes ad hoc

Eduardo Casilari Pérez, Alicia Triviño Cabrera
Departamento de Tecnología Electrónica. Universidad de Málaga
ETSI de Telecomunicación, Campus de Teatinos.
29071 - Málaga
Teléfono: 952 13 27 55 Fax: 952 13 14 47
E-mail: ecasilari@uma.es

***Abstract**. This work describes an empirical framework to define the simulated time that is required for a proper utilisation of mobility models in the simulation of an ad hoc network. In our analysis, link duration is utilised as a practical measurement of the effects of the mobility on the network performance. By means of thorough simulations with typical ad hoc scenarios which employ the Random Waypoint Model, the study tries to find and optimise the simulation time that guarantees a minimum skew in the estimation of link duration. The results confirm that fixing a common simulation time for all the possible mobility scenarios, as it is normally performed in the ad hoc literature, leads to strongly unstable and biased estimations of link duration. In order to solve this situation, the paper provides a practical worst-case formula to define the recommendable simulation time as a function of the parameters of the mobility model.*

1 Introducción

Las redes móviles ad hoc o MANET (*Mobile Ad hoc Network*) definen un nuevo paradigma para las redes de comunicaciones que desde hace ya más de un lustro viene concitando un vivo interés en la comunidad científica e industrial. Junto con todos los elementos que es necesario caracterizar en cualquier red de comunicaciones (desde los hábitos del usuario, a la carga y los modelos de tráfico o el comportamiento del medio físico,...), el estudio de las redes ad hoc exige la definición de modelos que describan la tipología e intensidad de la movilidad de los nodos que conforman la red. Esta movilidad es, en principio mucho más crítica que en otras tecnologías de comunicación móviles (como las de telefonía) ya que, en este caso, el alcance es menor y, además, la propia topología de la MANET se encuentra implícitamente definida por el movimiento de sus nodos.

Dada la complejidad de las variables que intervienen en una red *ad hoc*, el mecanismo de estudio y análisis de nuevas propuestas en este campo pasa a veces obligadamente por la simulación. En este campo la mayor parte de los estudios emplean masivamente el modelo de movilidad denominado *Random Waypoint*. Este modelo, cuya principal característica es su simplicidad, presenta no obstante una serie de problemas de estabilidad al ser empleado en simulaciones. En este artículo se define un marco que permite evaluar los requerimientos del uso de este modelo cuando se pretenden obtener resultados sin sesgo y estables en escenarios, donde, simultáneamente, se optimice el tiempo y la carga computacional de las simulaciones.

Este trabajo se estructura del siguiente modo: en el apartado 2 se define el modelo bajo estudio, comentando sus principales ventajas y desventajas. En la sección 3 se propone un marco de trabajo para estudiar el tiempo de simulaciones que empleen este modelo. El apartado 4 comenta las simulaciones efectuadas y los resultados obtenidos de ellas. Finalmente la sección 5 resume las conclusiones.

2 El modelo *Random Waypoint*

De acuerdo con el modelo de destino aleatorio o *Random Waypoint* (RWP) los nodos de una red ad hoc se desplazan en línea recta y con velocidad constante entre dos puntos elegidos al azar dentro del espacio limitado para los movimientos. Así, en el caso de un espacio bidimensional, para cada nuevo movimiento, el nodo determina las coordenadas (x e y) del siguiente destino mediante una variable aleatoria uniformemente distribuida entre 0 (origen de coordenadas) y el límite máximo permitido para los desplazamientos en cada dirección (x_{max} e y_{max}, respectivamente). Una vez que se alcanza un destino y previamente a elegir el siguiente, el modelo RWP permite pausas, normalmente caracterizadas con un valor constante (T_{pausa}). Igualmente, en la mayor parte de las implementaciones, la velocidad constante de cada trayecto se duele decidir también a través de una distribución uniforme en el intervalo (0, V_{max}], siendo V_{max} (expresada en m/s) la velocidad máxima permitida para los nodos.

Dejando a un lado los problemas del transitorio inicial que se comentarán más tarde, el modelo RWP padece una serie de inconvenientes que se desprenden de su propia simplicidad. Entre los más importantes podemos citar:

- Son escasos los escenarios reales en los que los nodos de comunicación se muevan de la manera "errática" o a la deriva que define RWP. El modelo ignora que tanto la velocidad como el destino de un nodo móvil no suelen ser procesos sin memoria y que, en cada trayecto, el valor de ambas variables se halla fuertemente determinado por el anterior.

- El modelo no contempla la presencia de obstáculos en el área de desplazamientos ni el hecho de que, en muchos escenarios de redes ad hoc, la movilidad se encuentra fuertemente determinada por la presencia de rutas prefijadas (calles, pasillos, senderos,...).

- El movimiento no considera posibles correlaciones en los movimientos de los nodos, los cuales se desplazan sin tener en cuenta a los demás. Esta circunstancia desprecia la evidencia de que en la mayoría de aplicaciones en donde el nodo es un agente humano, el movimiento sigue pautas fuertemente grupales (provocadas por la formación de pelotones, corrillos, unidades de salvamento...)

Para subsanar estas deficiencias la literatura ha propuesto (véanse, por ejemplo, [1], [6] o [8]) una serie de modelos que refinan los patrones de movilidad (incorporando correlaciones individuales o de grupo, definiendo la presencia de vías u obstáculos, etc) a costa de un incremento, a veces notable, de la complejidad y los parámetros requeridos por los procesos estocásticos involucrados.

En oposición a estos inconvenientes, el modelo RWP ofrece una enorme simplicidad y "parsimonia", esto es, una escasa necesidad de parametrización (limitada básicamente a los valores de T_{pausa}, V_{max} y V_{min}). Asimismo, al no estar enfocado a ningún escenario de aplicación concreto, su alto grado de abstracción lo convierte en un modelo generalista, idóneo para efectuar pruebas genéricas, al menos iniciales, ante nuevas propuestas de protocolos o mecanismos de gestión de recursos en redes MANET. Estas ventajas, junto con la facilidad y extensión de su implementación (el modelo se encuentra presente en las librerías de simuladores tan populares como NS [13]), seguramente han posibilitado que sea, con diferencia, el modelo más común en la literatura sobre redes ad hoc.

3. Marco y objetivos del estudio

3.1 Tiempo entre movimientos

Dada su amplia utilización, el estudio de las propiedades del modelo RWP como proceso estocástico ha provocado la aparición de diversos trabajos en los últimos años [2-5][7][9-12] [14-15]. En muchos de estos artículos se denuncia el empleo incorrecto que de este modelo normalmente efectúa la literatura. Efectivamente, si los parámetros de RWP y su propia ejecución no se programan adecuadamente, los resultados de una simulación con este modelo pueden verse seriamente afectados por problemas de transitorios e inestabilidad. Estos problemas se hacen tanto más evidentes si se tiene en cuenta que las simulaciones de redes ad hoc son muy costosas computacionalmente (se suele simular a nivel de paquete) y que, por esta razón, la mayoría de los estudios limitan a unas cuantas centenas de segundos el tiempo simulado de cada experimento. En esta línea la literatura anteriormente citada no ofrece recomendaciones prácticas que permitan al usuario del modelo RWP definir de un modo sencillo la temporación que conviene a sus simulaciones.

La mayoría de los trabajos (incluso buena parte de la bibliografía centrada en el modelo RWP) establecen el tiempo simulado (T_{sim}) de un modo arbitrario y sin tener en cuenta el escenario de movilidad emulado. Optimizar este tiempo implica elegir un valor mínimo que optimice los costes de computación pero que, al mismo tiempo, reduzca lo suficiente el sesgo que imponen en los resultados los problemas de transitorios e inestabilidad inherentes a RWP. Para estudiar este tiempo entendemos que es imprescindible considerar el grado de movilidad de los nodos. Para caracterizar dicha movilidad proponemos emplear el parámetro de tiempo entre movimientos (T_{mov}), definido como el tiempo medio esperado entre el inicio de dos movimientos consecutivos de cada nodo. Este tiempo resulta de la suma de dos componentes: el tiempo (en principio fijo) de pausa T_{pausa} y el tiempo medio invertido en cada desplazamiento.

De acuerdo con la fórmula en [11], la distancia media recorrida entre destino y destino por un nodo que se mueve siguiendo un modelo RWP se puede calcular como:

$$E(L)=\frac{1}{6}\left[\frac{y_{\max}^2}{x_{\max}}\log\left(\frac{\sqrt{\frac{y_{\max}^2}{x_{\max}^2}+1}+\frac{x_{\max}}{y_{\max}}}\right)+\frac{x_{\max}^2}{y_{\max}}\log\left(\frac{\sqrt{\frac{x_{\max}^2}{y_{\max}^2}+1}+\frac{y_{\max}}{x_{\max}}}\right)\right]+$$
$$+\frac{1}{15}\left(\frac{x_{\max}^3}{y_{\max}^2}+\frac{y_{\max}^3}{x_{\max}^2}\right)-\frac{1}{15}\sqrt{x_{\max}^2+y_{\max}^2}\left(\frac{x_{\max}^2}{y_{\max}^2}+\frac{y_{\max}^2}{x_{\max}^2}-3\right)$$

donde x_{max} e y_{max} representan las dimensiones del área de movimiento.

A partir de este valor, si se define $E(V)$ como la media esperada de la velocidad de los nodos, resulta evidente que el tiempo medio que invierten los nodos entre los inicios de dos movimientos es:

$$T_{mov}=\frac{E(L)}{E(V)}+T_{pausa} \qquad (1)$$

En [9] se demuestra que $E(V)$ no resulta del punto equidistante entre el valor mínimo (V_{min}) y el máximo (V_{max}), como inicialmente cabría esperar de una la velocidad generada mediante una distribución

uniforme. Por el contrario, su valor viene dado por la expresión: $$E(V) = \frac{V_{\max} - V_{\min}}{\ln\left(\frac{V_{\max}}{V_{\min}}\right)} \quad (2)$$

Este comportamiento se justifica por el hecho de que los nodos a los que el proceso aleatorio asigne velocidades más lentas tardarán más tiempo en detenerse y, por tanto, pesarán más a la hora de determinar la velocidad media.

De la expresión anterior se deduce la enorme importancia que presenta el valor de la velocidad mínima. Tal y como se recalca en [15], un valor nulo de este parámetro provoca que la velocidad media de los nodos tienda a converger a cero, provocando un transitorio infinito en las simulaciones que conduce, sea cual fuere el tiempo de simulación, a resultados fuertemente sesgados. Para evitar este efecto [15] propone definir una velocidad mínima. Si, añadidamente, se igualan los valores máximo y mínimo, fijando una velocidad constante para los nodos, este problema particular de convergencia se anula ya que es fácilmente demostrable que: $$\lim_{V_{\max} \to V_{\min}} (E(V)) = V_{\min} \quad (3)$$

Una vez definida V_{min}, resulta evidente que: $$E(V) > V_{\min} \quad \forall V_{\max} > V_{\min} \quad (4)$$

Por tanto, el caso de velocidad constante, aparte de reducir los parámetros del modelo RWP a dos (T_{pausa} y V_{min}) puede contemplarse como un caso "peor" o límite en lo que se refiere a la cadencia que la velocidad impone en la dinámica de las simulaciones.

En estas condiciones, el valor de esta cota para T_{mov} puede estimarse como: $$T_{mov} = \frac{E(L)}{V_{\min}} + T_{pausa} \quad (5)$$

3.2. Métricas de movilidad: duración del enlace

Para evaluar de un modo general la utilización del parámetro anterior T_{mov} como unidad del tiempo de simulación, es necesario definir una métrica que realmente caracterice la condiciones de movilidad de los enlaces en la red.

En la bibliografía se prueban diversos escenarios de movilidad asignando diferentes valores al tiempo de pausa y las velocidades del modelo RWP que rige los desplazamientos de los nodos. De este modo, es común que las prestaciones analizadas de la red se investiguen en función de alguna de estas variables (representadas en el eje x de los resultados gráficos). Sin embargo, uno sólo de estos parámetros no permite caracterizar por completo la variabilidad en la conectividad en una red ad hoc, ya que esta depende de la interacción entre ellos mismos (pausas y velocidades) y de su relación con otros elementos externos al propio modelo RWP como pueden ser el alcance radio de los enlaces o las propias dimensiones del área de simulación.

En [5] se comparan diversas métricas posibles para representar la movilidad de una red ad hoc, apostando finalmente por la duración del enlace como la más eficiente. Este estudio, aparte de probar la eficacia de esta métrica para estudiar las prestaciones de protocolos de encaminamiento, resalta otras ventajas añadidas de emplear esta variable como el hecho de que puede computarse en cada nodo sin necesidad de conocer la evolución global de la red, de que su estimación consume escasos recursos, de que es universal y no depende del protocolo que se esté empleando para encaminar, o de que se pueda medir de modo simple en escenarios reales.

A lo largo de la siguiente sección se estudia el valor necesario que ha de tener T_{sim} (en términos de T_{mov}) para definir simulaciones que arrojen valores insesgados y estables de esta métrica.

4 Simulaciones y resultados

Para analizar la influencia del tiempo de simulación sobre la estabilidad y el sesgo de las medidas de la duración del enlace, se consideraron diversas condiciones de movilidad en un entorno de simulación basado en Matlab. En la Tabla 1 se han incluido los valores que definen los escenarios de las simulaciones realizadas. Estos valores se encuentran en el rango típico de los que suele emplear comúnmente la literatura.

Tabla 1. Parámetros de los escenarios de simulación

Dimensiones	x=1500 m, y=300 m,($E(L)$=524.639 m)
Nº de nodos	50 (densidad=0.0001111 nodos/m^2)
Nº de ejecuciones por punto estimado	>5
Velocidad (V_{min}=V_{max})	1, 5 y 20 m/s
Tiempos de pausa	[0-50] s
Alcance	250 m
Distribución inicial de los nodos	Siguiendo su distribución estacionaria
Estado inicial de los nodos	Pausado o iniciando un desplazamiento de acuerdo con la probabilidad estacionaria de estar pausado

Como se puede observar en la tabla, básicamente se concretan tres tipos de pruebas para las que se define una velocidad constante y que podemos entender que describe un amplio abanico que engloba a la mayoría de las aplicaciones o escenarios de redes ad hoc con nodos móviles:

- Escenario de movilidad lenta: velocidad de 1 m/s (velocidad tipo de un ser humano andando)

- Escenario de movilidad media/rápida: velocidad de 5 m/s (caso de un ser humano corriendo o en un vehículo no motorizado).

- Escenario de movilidad muy rápida: velocidad de 20 m/s (asimilable a entornos en los que la movilidad se encuentra determinada por vehículos a motor).

Un campo de aplicación de las redes ad hoc de mucho interés que quedaría fuera de este análisis es el de las redes de sensores. En la mayoría de estas redes se puede entender que la movilidad es nula. Obviamente, por tanto, su estudio queda al margen de la problemática que puedan suscitar los propios modelos de movilidad.

En un primer intento de evaluar la importancia del tiempo de simulación, comparamos los tres escenarios anteriores, fijando tres tiempos de pausa de 0, 25 y 50 s y utilizando en todos los casos los mismos valores de T_{sim}. Las figuras 1, 2 y 3 muestran (para v_{min}=1 m/s, 5 m/s y 20 m/s respectivamente) las estimaciones de la media y la mediana de la duración del enlace para valores de T_{sim} elegidos en el rango de 50 a 4000 s (en el que habitualmente se mueven la inmensa mayoría de simulaciones con redes *ad hoc*). En todos las simulaciones, se entiende por duración del enlace aquel que transcurre entre que un nodo entra en el radio de cobertura de otro hasta que vuelve a escapar del mismo.

Del análisis de la figura se desprende la importancia de la velocidad en la convergencia de las estimaciones. Cotejando las tres figuras se observa claramente que velocidades menores obligan a tiempos de simulación mayores. Así, en el escenario muy rápido (v_{min}=20 m/s), un tiempo simulado de 1000 s conduce a un sesgo de la media inferior al 1% mientras que en el de movilidad menor el sesgo se encuentra en torno al 40% con respecto al valor alcanzado con un T_{sim} de 4000 s.

Otra conclusión del análisis de estas figuras, no por esperable menos importante, es que sólo se deben tener en cuenta escenarios con diversos tiempos de pausa cuando éstos son comparables al tiempo medio que el nodo emplea entre destino y destino. En otro caso, la definición de la pausa apenas provoca cambios en los resultados simulación. Si definimos como P_p la probabilidad o fracción de tiempo en la que un nodo se encuentra pausado, esta puede aproximar, con velocidad constante, mediante la expresión:

$$P_p = \frac{T_{pausa}}{T_{mov}} = \frac{T_{pausa}}{T_{pausa} + \frac{E(L)}{V_{\min}}} \qquad (6)$$

Si el cambio en el tiempo de pausa no provoca variaciones significativas de este parámetro, las simulaciones así definidas no añaden nada al escenario investigado. Por ejemplo, para la velocidad menor (v_{min}=1 m/s), las pausas de 0, 25 y 50 s apenas provocan que el valor de esta probabilidad alcance un 10% (P_p=0, 0.045 y 0.087 respectivamente), lo que se refleja en que las estimaciones de la duración del enlace prácticamente coincidan. Esto quiere decir, a efectos prácticos, que en escenarios bidimensionales típicos (con dimensiones de centenas de metros), con baja velocidad de los nodos (menores de 2 m/s), pausas inferiores a un minuto apenas tienen efecto. Se propone, por tanto, que en simulaciones de redes *ad hoc* la variación del tiempo de pausa se establezca en función de cambios en este parámetro P_p y no con incrementos absolutos de T_{pausa} independientes de la velocidad mínima definida.

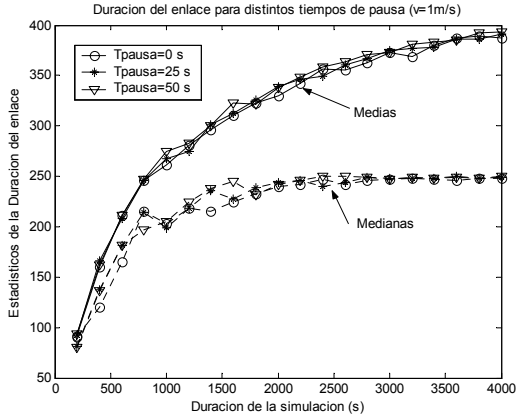


Figura 1. Estadísticos de la duración del enlace (velocidad=1 m/s)

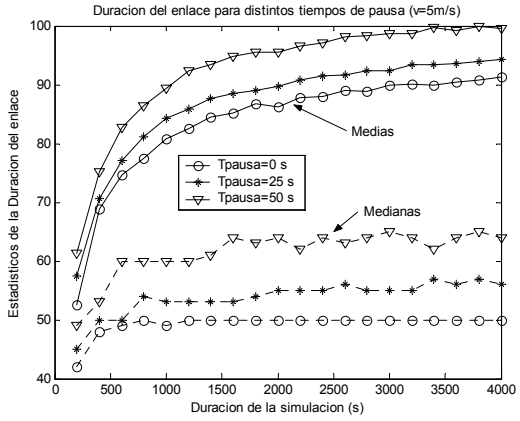


Figura 2. Estadísticos de la duración del enlace (velocidad=5 m/s)

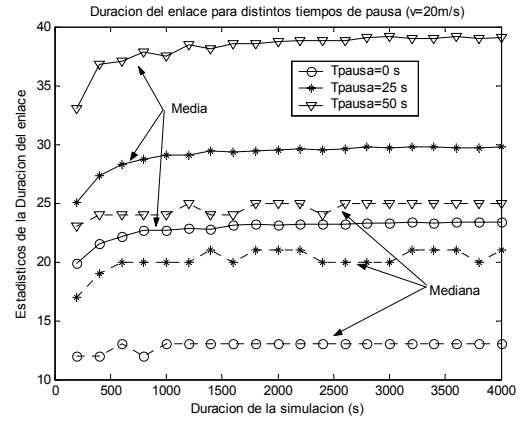


Figura 3. Estadísticos de la duración del enlace (velocidad=20 m/s)

En las figuras anteriores se observa claramente que la convergencia de la mediana a su valor final es mucho más rápida que el de la media. Por tanto, en simulaciones cortas el empleo de este estadístico resultaría mucho más adecuado que el de la media. Esto se debe a que el estimador de la mediana no se ve afectado por los valores más altos en las estimaciones de la duración del enlace, los cuales aumentan conforme crece el tiempo de simulación. La media, por el contrario, es sumamente sensible a la "cola" de la distribución de la variable que esté representando. Para estudiar la forma de dicha distribución, se efectuaron simulaciones muy largas con los tres escenarios (de 1000 veces el valor de T_{mov}) y se estimó la función de distribución complementaria o CDF (*Complementary Distribution Function*). La figura 4 muestra que la caída de la cola de la distribución presenta una forma exponencial (lineal con el eje de ordenada en escala logarítmica) con independencia de la velocidad empleada. No obstante, la figura, que incluye distintos ajustes con distribuciones estándares, también ilustra cómo un puro ajuste exponencial es incapaz de aproximar correctamente esta caída. En concreto el análisis indica un mejor ajuste mediante la distribución de Weibull, la cual es más flexible a la hora de caracterizar caídas exponenciales, ya que su definición permite un grado más de libertad que la función exponencial. En concreto su función de densidad sigue la expresión:

$$f_W(x) = \frac{b}{a^b} \cdot \exp\left[(-\frac{x}{a})^b \right] \cdot x^{b-1} \quad if \quad 0 \le x \le \infty \quad (7)$$

siendo a y b los factores de escala y forma, respectivamente. La distribución exponencial equivale a una distribución de Weibull con factor de forma unidad (b=1) mientras que el ajuste lineal que propone una regresión de mínimos cuadrados de la caída de la cola arroja valores del factor de forma muy similares para las tres curvas y siempre menores que la unidad. De acuerdo con esto y con las caídas representadas en la figura 4, el factor de forma parece probarse independiente de la velocidad de los nodos la cual, por tanto, sólo afecta al factor de escala (a mayor velocidad, menor duración).

Esta distribución con una caída más lenta (b<1) que la puramente exponencial de las duraciones de los enlaces contradice el modelo propuesto en [10]. En dicho trabajo, en lugar de emplear un modelo de movilidad, se propone caracterizar la posible conectividad entre dos nodos mediante un modelo de tipo On Off. Esta estructura markoviana presupone que la permanencia en cada estado sigue una distribución exponencial, algo que, como se ha probado, no se cumple en un régimen de movilidad que emplee *Random Waypoint*.

En cualquier caso, la variabilidad de la distribución se reduce al introducir en los patrones de movilidad pausas fijas. En la figura 5 muestra cómo, para la misma velocidad (v=20 m/s), el incremento en la pausa provoca una mayor rapidez en la caída de la distribución (con factores de forma que superan la unidad). Lógicamente esto se explica por el hecho de que conforme T_{pausa} aumenta (y P_p se acerca a 1) el tiempo de pausa resulta mucho más determinante en la duración del enlace que el que invierten los nodos al desplazarse y, por tanto, la variabilidad relativa se ve limitada por este valor constante.

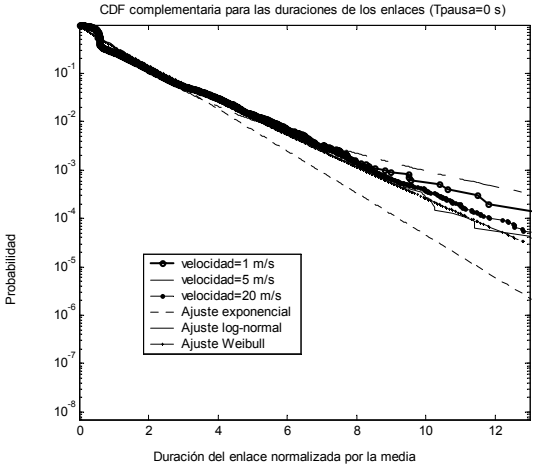


Figura 4. CDF de la duración de enlace (T_{pausa}=0)

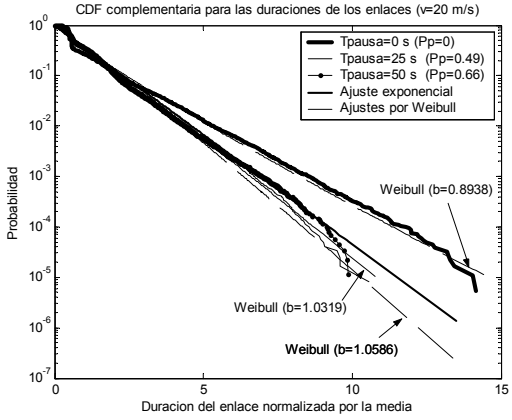


Figura 5. CDF de la duración de enlace (velocidad=20 m/s)

De los análisis anteriores se deduce que la velocidad mínima y los tiempos de pausas determinan el rango de tiempos en el que se mueve la dinámica de la simulación. Por tanto resulta claramente incorrecto fijar T_{sim} sin tener en cuenta estos valores de V_{min} y T_{pausa}. Para tratar de analizar esta influencia se repitieron las simulaciones pero, en esta ocasión, el tiempo de simulación se estableció en proporción al tiempo medio entre movimientos (T_{mov}) de cada escenario.

La figura 6 muestra, para los tres escenarios con pausas nulas, la estimación de la media y la mediana de la duración del enlace en función del parámetro de proporción r, definido como la relación entre los tiempos T_{sim} y T_{mov}.

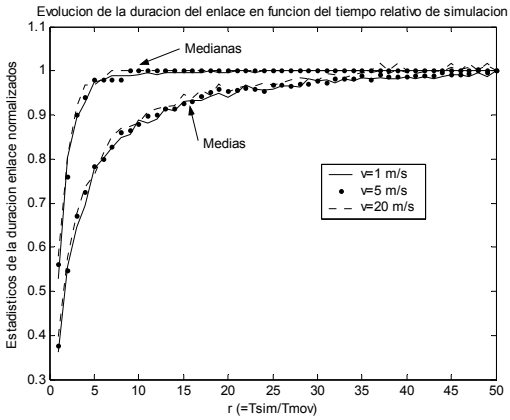


Figura 6. Estadísticos de la duración del enlace (T_{pausa}=0 s)

Los resultados, normalizados por los valores de la media respectiva, no muestran ninguna divergencia significativa entre los tres escenarios. En consecuencia, si se establece T_{sim} fijando un valor para el parámetro r para todas las simulaciones, el sesgo de las medidas no dependerá del grado de movilidad de los nodos. Añadidamente, a la vista de los resultados obtenidos, se comprueba que basta un valor de r =5 para conseguir un valor estable de la mediana de la duración del enlace mientras que la estabilidad de la media (alcanzando un sesgo inferiores al 1%) exige un r en torno a 50.

En cuanto al intervalo de confianza de la propia medida de la media, también se prueba de las simulaciones obtenidas (véase la figura 7) que queda reducido a valores por debajo del 1% (fijando la confianza en el 99% y asumiendo la distribución de *t-student* con 9 grados de libertad en las estimaciones) cuando r toma un valor próximo a 50.

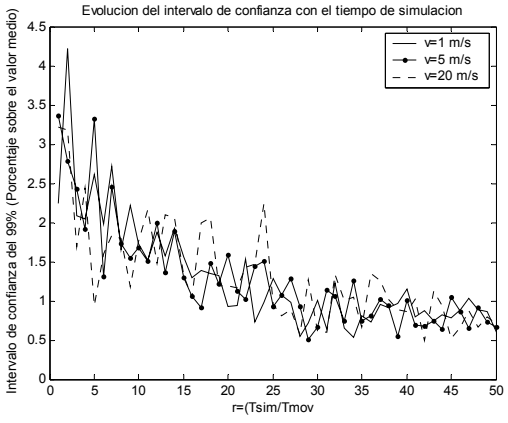


Figura 7. Evolución del intervalo normalizado de confianza (al 99%) de la medida de la duración media

El caso de pausa nula puede considerarse en este ámbito como un caso peor. Como ya se mostró, pausas mayores reducen la variabilidad de las duraciones y, por tanto, el sesgo en los estimadores. Esto se puede observar en la figura 8 donde se representa la evolución del estimador de la media en función de la probabilidad P_p para distintos valores de la proporción r. La figura incluye el caso de una simulación muy larga (r=1000) mostrando que sus resultados apenas difieren de los obtenidos con r=50, incrementando por tanto, innecesariamente el coste computacional. Asimismo, se observa que al aumentar P_p los resultados de las diversas gráficas convergen, reduciéndose la necesidad de un tiempo mayor de simulación para estimar la duración media.

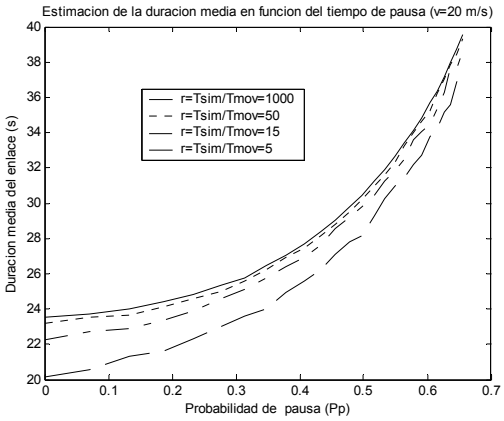


Figura 8. Estimación de la duración media del enlace para distintos valores del tiempo de pausa (velocidad=20 m/s)

4.1. Efecto del transistorio con distribuciones iniciales uniformes

Existen numerosos estudios en la literatura [2] [3] [4] [7] [11] [14] que prueban que las distribuciones estacionarias de la posición y velocidad de los nodos no siguen un modelo uniforme debido al propio tiempo de tránsito. Esto quiere decir que si los valores inicialmente elegidos para estas variables siguen una distribución uniforme, existirá un transitorio durante el que la distribución de ambas variable tenderá a su valor estacionario.

Para evitar esta distribución estacionaria no uniforme hay estudios como [4] que proponen utilizar una variante del modelo RWP que elige los puntos de destino de los trayectos justo en las fronteras del área de simulación. En [7] también se estudia analíticamente esta variación, a la que se denomina RWPB (*Random Waypoint on the Border*).

Una solución que no obliga a redefinir el modelo es la propuesta en [11] y [12]. Estos trabajos proponen elegir directamente los valores de velocidad y posición de los nodos de acuerdo con su distribución estacionaria en lugar de emplear una uniforme. Con ello se logra eliminar o reducir en gran medida el efecto del transitorio ya comentado. Esta estrategia, tal y como se comenta en la tabla 1, es la que se ha seguido en las simulaciones presentadas hasta ahora. No obstante, dado que la mayoría de los estudios no tienen en consideración la existencia de este transitorio, cabe preguntarse cuál es su influencia en la estimación de la movilidad de la red. Para evaluar este aspecto se plantea repetir las simulaciones,

empleando una distribución inicial de los nodos generada de acuerdo con una distribución uniforme y comparar sus resultados con los anteriores (el problema de la convergencia de la velocidad no existe en tanto que esta se considera constante).

La figura 9 muestra esta comparativa para el caso del escenario de movilidad baja, probando que el transitorio apenas afecta a la medición de la duración de los enlaces transcurrido el tiempo necesario para evitar el sesgo de las medidas.

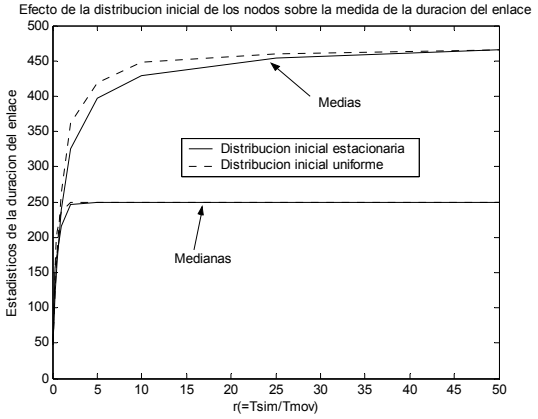


Figura 9. Estadísticos de la duración del enlace con distribuciones iniciales de los nodos uniforme y estacionaria (T_{pausa}=0)

Donde cabe esperar que la distribución inicial posea más relevancia es en el estudio del número de enlaces activos. Efectivamente, el desplazamiento de los nodos con RWP tiende a elegir como paso la zona central del área de simulación, razón por la que en condiciones estacionarias los nodos se encuentran más concentrados que con la distribución uniforme. Por tanto, conforme la simulación avanza, si se partió distribuyendo homogéneamente los nodos por el área, el número de enlaces activos aumenta. Este efecto se ilustra en la figura 10. Los resultados de esta figura muestran, al mismo tiempo, que, con independencia de la velocidad empleada, el transitorio se puede dar por finalizado tras un periodo equivalente a un tiempo entre movimientos (r=1). Otra manera de justificar la brevedad del estacionario es estudiar la distribución de los nodos para distintos tiempos de simulación, tal y como se hace en la figura 11. La figura también incluye la representación de la distribución estacionaria ($F(x)$) que se debe alcanzar. Dado que la función de densidad espacial en la dirección x, en régimen estacionario, de un nodo que se mueve mediante un proceso RWP, es [14]:

$$f(x) = P_p + (1-P_p)6\left[\frac{x}{x_{\max}} - \left(\frac{x}{x_{\max}}\right)^2\right] \quad x \in [0, x_{\max}] \qquad (8)$$

integrando, se obtiene:

$$F(x) = P_p x + (1-P_p)\left[3\frac{x^2}{x_{\max}} - 2\left(\frac{x}{x_{\max}}\right)^3\right] \quad x \in [0, x_{\max}] \qquad (9)$$

Los resultados muestran que, tras un tiempo de simulación igual al de movimiento de los nodos (r=1), la distribución de los nodos (aun partiendo de un valor uniforme) alcanza un valor muy cercano al estacionario.

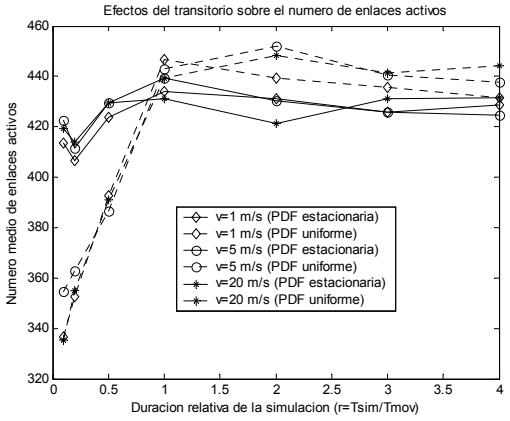


Figura 10. Estimación del número de enlaces con distribuciones iniciales de los nodos uniforme y estacionaria (T_{pausa}=0)

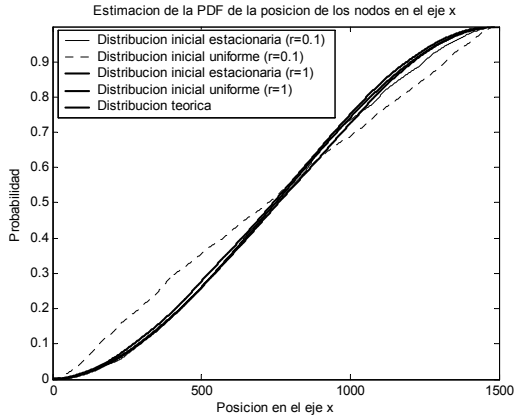


Figura 11. Distribución estimada de la posición de los nodos en el eje x para diversos tiempos de simulación (v_{min}=1m/s,T_{pausa}=0 s)

De las pruebas anteriores se deduce que el tiempo del transitorio provocado por distribuciones iniciales uniformes es despreciable frente al que se necesita para obtener estadísticos insesgados de la duración del enlace.

5 Conclusiones

En este trabajo se ha estudiado mediante simulaciones sistemáticas la problemática del modelo de movilidad *Random Waypoint* para redes ad hoc. En particular el análisis se ha enfocado para determinar una regla empírica que permita elegir un tiempo de simulación óptimo. En esa línea se prueba que la rapidez para alcanzar el régimen estacionario (entendiendo como tal aquel en el que se estabilizan los estadísticos relativos a la duración de los enlaces) depende fuertemente de la velocidad de los nodos. Del estudio efectuado se deduce que el tiempo simulado necesario para alcanzar resultados estables es directamente proporcional al tiempo medio que un

nodo invierte entre inicio e inicio de dos movimientos consecutivos. En consecuencia, resulta inadecuado el procedimiento comúnmente empleado en la literatura de utilizar el mismo tiempo de simulación para probar diversos escenarios de simulación.

Basándonos en los resultados empíricos obtenidos, se puede afirmar que, para alcanzar resultados con un sesgo y un margen de confianza inferiores al 1%, el tiempo de una simulación ha de cumplir:

$$\frac{T_{sim}}{T_{mov}} = \frac{T_{sim}}{\frac{E(L)}{V_{\min}} + T_{pausa}} \geq r_{\min} \qquad (10)$$

donde *E(L)* depende de las dimensiones del espacio de y se calcula a partir de la fórmula de la sección 3.1 y r_{min} es una cota mínima. De las simulaciones efectuadas se deduce que si la estimación se centra en la mediana de la duración basta un r_{min} de 5 mientras que si lo que se estima es la media se precisa un valor para r_{min} de 50. Estas cotas mínimas aunque no necesariamente las más restrictivas (ya que no consideran los transitorios impuestos por las fuentes de tráfico, aspecto también crucial en toda simulación) pueden ser extensibles a cualquier simulación de una red hoc para la que se defina una velocidad mínima. Además resultan independientes del número de nodos que se definan ya que la duración de los enlaces no depende de esta variable. Por otro lado, como consecuencia importante de la fórmula anterior, y tal y como se apunta en [15], resulta clave el valor con que se defina esta velocidad mínima. Normalmente, la literatura no restringe este valor en la medida en que la velocidad de los nodos se elige uniformemente en el intervalo (0, V_{max}]. Sin embargo, conforme con los resultados presentados, establecer valores muy bajos para esta variable puede obligar a simulaciones de una duración inaceptable desde un punto de vista práctico. Así se propone fijar una velocidad mínima nunca inferior valores de 1 m/s. Entendemos que este valor caracteriza razonablemente la movilidad mínima de la mayoría de los escenarios de aplicación de redes ad hoc.

Agradecimientos

Este trabajo ha sido parcialmente costeado por el proyecto Nº TEL2003-07953-C02-01.

Referencias

[1] F. Bai, N. Sadagopan, A. Helmy, "The IMPORTANT Framework for Analyzing the Impact of Mobility on Performance of Routing for Ad Hoc Networks", *AdHoc Networks Journal*, Vol. 1, Issue 4, pp.383-403, Nov. 2003.

[2] C. Bettstetter, G. Resta, P. Santi, "The Node Distribution of the Random Waypoint Mobility Model for Wireless Ad Hoc Networks", *IEEE Transactions on Mobile Computing*, Vol. 2, No. 3, pp 257-269, July-Sept 2003.

[3] C. Bettstetter, H. Hartenstein, X. Pérez-Costa, "Stochastic Properties of the Random Waypoint Mobility Model", *ACM/Kluwer Wireless Networks,* vol. 10, no.5, Sept 2004, pp. 555-567.

[4] C. Bettstetter, C. Wagner, "The Spatial Node Distribution of the Random Waypoint Mobility Model", *GI Lecture Notes in Informatics*, no.P-11, pp. 41-58, 2002.

[5] J. Boleng, W. Navidi, T. Camp, "Metrics to Enable Adaptive Protocols for Mobile Ad Hoc Networks", *Proc. of ICWN '02*, Las Vegas, Nevada, USA, pp. 293-298, Jun 24-27, 2002.

[6] T. Camp, J. Boleng, V. Davies, "A Survey of Mobility Models for Ad Hoc Network Research", *Wireless Communication & Mobile Computing,* Vol. 2, No. 5, pp. 483-502, 2002

[7] E. Hyytiä, P. Lassila, L. Nieminen, J. Virtamo, "Spatial Node Distribution in the Random Waypoint Mobility Model", Informe técnico, Helsinki University of Technology, 2004.

[8] A. Jardosh, E. M. Belding-Royer, K. C. Almeroth, and S. Suri. "Towards Realistic Mobility Models for Mobile Ad hoc Networks." *Proc. of MobiCom*, San Diego, CA, Sept. 2003.

[9] G. L. Lin, G. Noubir, and R. Rajaraman, "Mobility Model for Ad Hoc Network Simulation," *Proc. of IEEE INFOCOM 2004*, Hong Kong (China), 2004.

[10] T. Lin and S. F. Midkiff, "Mobility versus Link Stability in the Simulation of Mobile Ad Hoc Networks", Proc. of CNDS, pp. 3-8, Orlando (FL, USA) Jan. 2003.

[11] W. Navidi, T. Camp, "Stationary Distributions for the Random Waypoint Mobility Model", *IEEE Transactions on Mobile Computing*, vol. 3, no.1, pp.99-108, Jan-March 2004.

[12] W. Navidi, T. Camp, and N. Bauer, "Improving the Accuracy of Random Waypoint Simulations Through Steady-State Initialization", *Proc. of the 15th International Conference on Modeling and Simulation*, pp. 319-326, March 2004.

[13] Network Simulator, ns-2. Software disponible en: http://www.isi.edu/nsnam/ns/

[14] G. Resta, P. Santi, "An Analysis of the Node Spatial Distribution of the Random Waypoint Model for Ad Hoc Networks", *Proc. of POMC*, Toulouse (France), October 2002, pp. 44-50.

[15] J. Yoon, M. Liu, and B. Noble, "Random Waypoint Considered Harmful", *Proc. of IEEE INFOCOM 2003*, San Francisco, 2003.

Evaluación Experimental de las prestaciones de la Tecnología Bluetooth

V. Téllez García-Moreno, J. I Moreno Novella, A. Cuevas Casado, P. A. Vico Solano, D.Haage
Departamento de Ingeniería Telemática. Universidad Carlos III de Madrid
Av. De la Universidad, 30. Edif. Torres Quevedo.
28911 Leganés (Madrid)
Teléfono: (+34) 91 624 8778
E-mail: {vtellez@inv.uc3m.es, joseignacio.moreno, antonio.cuevas, pedroantonio.vico}@it.uc3m.es,

***Abstract**. The Bluetooth technology is a global specification, which provides short range, low power, low cost, pico-cellular wireless connectivity. In this paper we evaluate Bluetooth capabilities in different scenarios, and measure performance in terms of throughput and access delay. We use a Bluetooth platform and perform different measurements over selected scenarios to try to get results for helping the dimensioning and the engineering task.*

1.-Introducción

En los últimos años, los servicios habilitados por las tecnologías inalámbricas han permitido un importante desarrollo de las telecomunicaciones y una gran aceptación en la sociedad actual.

Bluetooth, que originalmente fue pensada para facilitar la sustitución de cables en entornos ofimáticos/personales, es una tecnología cuyo principal objetivo es facilitar la comunicación entre equipos móviles y fijos, eliminando cables y conectores posibilitando la transmisión de voz y datos entre diferentes equipos mediante un enlace radio en la banda de 2.4 GHz, libre de cánones y reservada para propósitos médicos, científicos o industriales.

Este artículo tiene el objetivo de mostrar los resultados de experimentos realizados con el fin de determinar las capacidades de la tecnología Bluetooth. Para ello se han desarrollado medidas de distintos parámetros en escenarios reales. Se determinará si, como se pensó en un principio, tan sólo se puede utilizar Bluetooth para reemplazar cables o si por el contrario el potencial de la tecnología es mayor, ofreciéndonos una plataforma barata y sencilla para crear nuevos servicios.

Los resultados derivados de estos experimentos podrán servir como base para determinar los campos de aplicación de la tecnología, los factores que influyen en su comportamiento y limitaciones de la misma, pudiendo utilizar las medidas como referencia a la hora de dimensionar e implementar servicios con Bluetooth.

Existen numerosos estudios que han abordado este tema realizando un análisis de las prestaciones, evaluando el comportamiento de Bluetooth en distintos entornos o bien su coexistencia con otras tecnologías que utilizan la misma banda del espectro [1,2,3,4,5,6,7]. En la mayoría de los casos este estudio se realiza desde un punto de vista teórico o utilizando simulaciones.

Con nuestro estudio queremos ir más allá aportando datos empíricos y determinando de forma real las posibles trabas con las que pueden encontrarse los desarrolladores de servicios al utilizar esta tecnología.

En el siguiente apartado realizaremos un breve resumen de las características más importantes del estándar Bluetooth.

El tercer apartado esta dedicado a describir el procedimiento de obtención de medidas, así como los parámetros, entornos y herramientas de desarrollo utilizadas. Una vez descrita la metodología se explicaran los resultados de las distintas pruebas realizadas, terminando con un apartado de conclusiones donde se resumirán los aspectos mas destacados de la evaluación.

2.-Estándar Bluetooth

Bluetooth es un estándar para comunicaciones radio de corto alcance y redes de área personal (PAN) desarrollado por el SIG ("Bluetooth Special Interest Group") [8,9]. El estándar Bluetooth utiliza la banda de 2.4 GHz, reservada internacionalmente para propósitos científicos y médicos, emplea técnicas de espectro ensanchado con salto de frecuencia (FHSS) para evitar interferencias. El canal está representado por una secuencia pseudoaleatoria de canales de radiofrecuencia. Para la comunicación se utiliza un esquema TDD ("Time Division Duplex") donde maestro y esclavo transmiten alternativamente. Cada slot corresponde con una frecuencia de salto distinta con una duración de 625 μsg, realizándose 1600 saltos por segundo.

Bluetooth proporciona una conexión punto a punto o punto a multipunto. Dos o más dispositivos que comparten un canal forman una *piconet*. Un dispositivo dentro de la *piconet* actuará como maestro de la *piconet*, mientras que el resto lo harán como esclavos. Como máximo pueden estar activos siete esclavos en una *piconet*.

Múltiples *piconets* con áreas de cobertura solapadas forman una *scatternet*. Un dispositivo puede participar en varias *piconets* simultáneamente.

Entre maestro y esclavo se pueden establecer dos tipos de conexiones: enlace orientado a conexión síncrono (SCO) y enlace no orientado a conexión asíncrono (ACL).

El enlace SCO es simétrico, está formado por una conexión punto a punto entre el maestro y un esclavo con un ancho de banda constante, reservado por el maestro para el intercambio de datos, el maestro puede soportar hasta tres enlaces SCO con el mismo o diferentes esclavos. El enlace ACL es un enlace punto-(multi)punto con tasa binaria no constante entre el maestro y esclavo.

Bluetooth define distintos tipos de paquetes relacionados con el tipo enlace físico utilizado. Como en nuestro estudio nos hemos centrado en el transporte de datos tan sólo consideraremos los paquetes ACL.

En Bluetooth se definen seis tipos de paquetes ACL, que se diferencian por su tamaño y la utilización de códigos de protección de errores. Los paquetes DM utilizan un código de protección FEC (Forward Error Correction) 2/3, es decir cada dos bits se añade uno de redundancia. En estos paquetes a cada 10 bits de datos se le añaden 5 bits de paridad, sin embargo los paquetes DH carecen de esta redundancia. Además, dependiendo de si el paquete ocupa 1, 3 o 5 slots acompañaremos al tipo de paquete con un número que indica el número de ranuras que ocupa. En la tabla 1 se resumen el tipo de paquetes ACL y sus principales características.

Al diseñar la arquitectura de protocolos se intentó maximizar el número de aplicaciones que pudieran implementarse e interactuar con Bluetooth, de forma que sobre los protocolos propios de Bluetooth se puedan reutilizar los existentes. En la figura 1 se muestra la capa de protocolos de Bluetooth.

Tipo	Datos de usuario (bytes)	FEC	Tasa Asimética (Kbps)
DM1	0-17	2/3	108.8
DH1	0-27	No	172.8
DM3	0-121	2/3	387.2
DH3	0-183	No	585.6
DM5	0-224	2/3	477.8
DH5	0-339	No	732.2

Tabla 1- Paquetes ACL

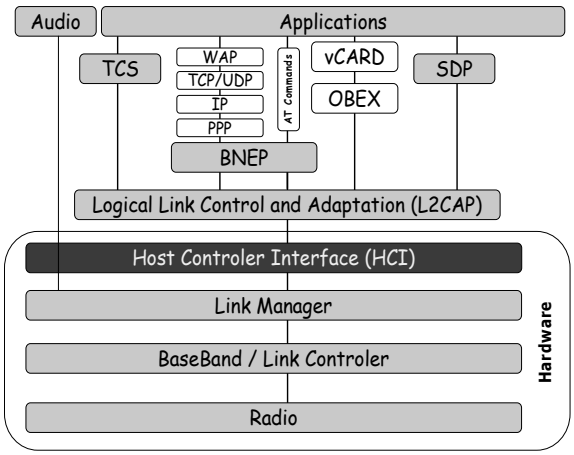


Figura 1-Pila de protocolos Bluetooth

3.- Configuración de las pruebas

Un punto importante a precisar fue la elección de una metodología de pruebas: elegir unos parámetros de medida que fueran lo suficientemente representativos para tener un conocimiento de las capacidades de la tecnología, así como los escenarios dónde evaluar esos parámetros.

Hay que resaltar que la lectura y comprensión de estudios relacionados con Bluetooth nos ha servido de gran ayuda. Aunque en la mayoría de los artículos se estudian las prestaciones basándose en estudios teóricos o simulaciones, nos han aportado muchas ideas para determinar los factores que influyen en las prestaciones de una tecnología inalámbrica como es Bluetooth. Destacamos los desarrollados por el grupo de trabajo TG2 de la IEEE 802.15 [10], donde además se proponen unas pautas de evaluación de la tecnología seguidas en la realización de nuestros experimentos [11].

Además de realizar las pruebas, fue preciso realizar un estudio de los posibles dispositivos Bluetooth que hay disponibles en el mercado y de las plataformas de desarrollo, eligiendo las más idóneas para la realización de las pruebas. Como plataforma de desarrollo se eligió finalmente BlueZ [12] por su flexibilidad y facilidad de uso.

En los siguientes apartados se describen de forma concisa los parámetros y escenarios de prueba elegidos, así como la pila de protocolos y los equipos Bluetooth utilizados en nuestros experimentos.

3.1-Escenarios de pruebas

A la hora de definir los escenarios dónde se han realizado las pruebas, se tuvo un especial cuidado en elegir entornos bien diferenciados, de forma que la introducción de cada escenario incluyese un factor que pudiese alterar las condiciones de propagación de la señal radio a la vez que resultaran escenarios de interés en el entorno de aplicación de esta tecnología.

Hay que tener en cuenta que los escenarios utilizados en las pruebas bien podrían ser escenarios reales, por lo que los resultados de los experimentos y las conclusiones aportadas por los mismos nos servirán para dimensionar o determinar la posibilidad de implementar un servicio específico sobre Bluetooth.

Los cuatro escenarios utilizados para realizar las pruebas son los siguientes:

- *Primer Escenario*: Entorno libre con línea directa de visión entre los dispositivos transmisor/receptor. Como localización real de este escenario se eligió una superficie abierta de dimensiones 11 m x 33 m.
- *Segundo Escenario:* Entorno cerrado con línea directa de visión entre transmisor y receptor. Las pruebas en este escenario se realizaron en un pasillo de dimensiones 3,20m x 40 m, libre de obstáculos que entorpezcan la comunicación.
- *Tercer Escenario:* Entorno cerrado con obstáculos de distinta naturaleza. Trata de simular un ambiente ofimático y evaluar el impacto que los obstáculos pueden tener en las comunicaciones de los dispositivos Bluetooth. Los resultados obtenidos en este escenario son muy relevantes por ser el ámbito natural de aplicación de esta tecnología.
- *Cuarto Escenario:* Entorno cerrado con interferencias provocadas por un par de dispositivos Bluetooth. Seleccionado para determinar la influencia que pueden tener en las comunicaciones entre dispositivos la proximidad de otros de la misma naturaleza.

3.2-Parámetros evaluados

- *Tasa binaria a nivel de enlace*

En los anteriores escenarios se medirá la tasa binaria a nivel de enlace en función de la distancia, utilizando los distintos tipos de paquetes que pone a nuestro alcance la tecnología Bluetooth.

Para realizar la medida de este parámetro se desarrollo una aplicación cliente/servidor con BlueZ, de forma que el dispositivo utilizado para medir la tasa binaria se encuentra en un bucle de espera hasta que se recibe una conexión. La tasa binaria resultante se obtiene como media de las 200 muestras obtenidas al llenar un buffer de datos de 2000 bytes en el equipo receptor, comprobando que para una distancia de 5 m el intervalo de confianza del 99% es menor del 5% del valor medio de las muestras tomadas.

- *Tasa binaria a nivel TCP*

Dada la extensión que tienen hoy en día los protocolos TCP/IP es interesante evaluar su comportamiento sobre una tecnología inalámbrica. Mediante estas pruebas no sólo se pretende evaluar de forma cuantitativa la tasa binaria a este nivel, sino que además se quería comprobar la viabilidad de implementar aplicaciones sobre Bluetooth que utilicen este protocolo.

TCP ofrece una comunicación extremo a extremo orientada a conexión y un servicio de transporte fiable a nivel de aplicación con control de flujo y congestión. Fue diseñado para transmitir datos a través de un enlace cableado, donde la probabilidad de error es muy pequeña y la pérdida de un paquete se puede considerar fruto de la congestión, de forma que cuando se detecta un fallo en la secuencia se supone que la red está saturada, disminuyendo la tasa de envío de paquetes hasta que la situación mejora. Debido a esta suposición, el empleo del protocolo TCP tiene un comportamiento más bien pobre cuando se utiliza en enlaces inalámbricos, donde la principal causa de las pérdidas de paquetes no se deben a la congestión, sino a la pérdida de paquetes debido a las condiciones del canal de comunicaciones.

Dada la importancia que tiene la implantación de estos protocolos sobre Bluetooth, hay muchos artículos en donde se trata esta problemática. Y se proponen algoritmos para solucionar el efecto de esta disminución de rendimiento [13, 14, 15, 16].

Para realizar las medidas se ha utilizado la capa BNEP (*Bluetooth Network Encapsulation Protocol*), encargada de encapsular los datos del protocolo IP [17]. En la figura 2 se muestra la pila de protocolos utilizada en estas pruebas.

Para medir el ancho de banda utilizaremos un software de licencia pública "netperf" [18] que emplea un programa cliente y otro servidor para monitorizar el intercambio de paquetes durante 10 segundos midiendo el ancho de banda medio.

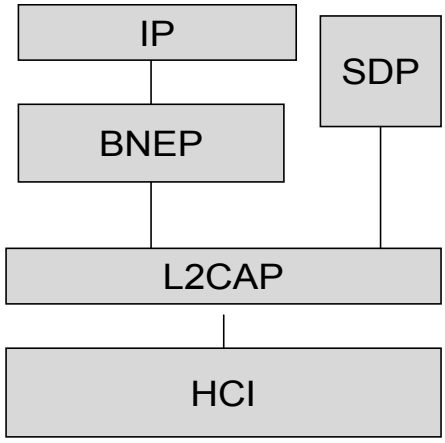


Figura 2-Estructura de protocolos utilizada para implementar TCP/IP sobre Bluetooth

- *Tiempo medio de acceso*

En redes inalámbricas dónde la creación de redes se hace de forma dinámica, dependiendo de los dispositivos que existan en las inmediaciones y de los servicios que ofrezcan. En este tipo de redes un parámetro muy importante es el tiempo medio de acceso, entendiéndolo como el tiempo transcurrido desde que un dispositivo intenta ponerse en contacto con otro hasta que se ha establecido un enlace entre ellos.

El tiempo medio de acceso es crucial en sistemas dinámicos o en aplicaciones dependientes de contexto, dónde los dispositivos actuarán de una determinada forma dependiendo de su situación, puesto que este tiempo medirá la adaptabilidad del sistema y el tiempo de acceso al servicio.

Desde el punto de ahorro energético en dispositivos móviles, también es importante este parámetro, al ser los procesos de búsqueda y conexión los que más potencia consumen.

Siguiendo el estándar Bluetooth para establecer una conexión entre dos dispositivos son necesarias dos fases: en la primera fase "inquiry", el dispositivo que se quiere incorporar envía mensajes para obtener los parámetros necesarios para establecer comunicación. Una vez conocidos éstos, se inicia la segunda fase "page" y se establece la comunicación.

Para realizar las medidas del tiempo de acceso se ha desarrollado un programa utilizando los comandos que pone a nuestro alcance BlueZ para comunicarnos con el nivel HCI ("Host Controller Interface") del dispositivo físico.

En el segundo escenario, se han realizado medidas del tiempo medio de acceso a cuatro distancias distintas para determinar la evolución de este parámetro con la distancia. Para cada distancia, se han tomado 100 muestras del tiempo de acceso, comprobando *a posteriori* que el número de muestras tomado es suficiente para que el intervalo de confianza al 95% sea de un 10% de variación.

3.3-Plataforma de desarrollo utilizada

Se realizó un estudio sobre las posibles pilas de protocolos que se pueden utilizar con Bluetooth. Existen en el mercado muchas pilas de protocolos (AXIS [19], Affix [20], BlueZ [12]), implementadas en diversos lenguajes de programación y sobre distintas plataformas.

Para el desarrollo de nuestras pruebas hemos elegido BlueZ por su flexibilidad, arquitectura modular y facilidad de uso. Además, el que sea la pila de protocolos oficial de Linux hace que esté en continuo desarrollo y tenga un gran apoyo entre los desarrolladores de todo el mundo. En la figura 3, se muestra la pila de protocolos de BlueZ.

3.4-Hardware y software utilizado

Para la realización de las medidas se utilizaron cuatro módems Bluetooth USB 3COM modelo 3CREB96B, al ser dispositivos de tipo 2 transmiten una potencia máxima de 4 dBm, lo que supone una cobertura teórica de 10m..

Se instaló la pila de protocolos de BlueZ sobre dos ordenadores: el nodo fijo utiliza un sistema operativo RedHat-7.1 con versión del kernel 2.4.15 mientras que el nodo móvil utiliza Mandrake 9.1.

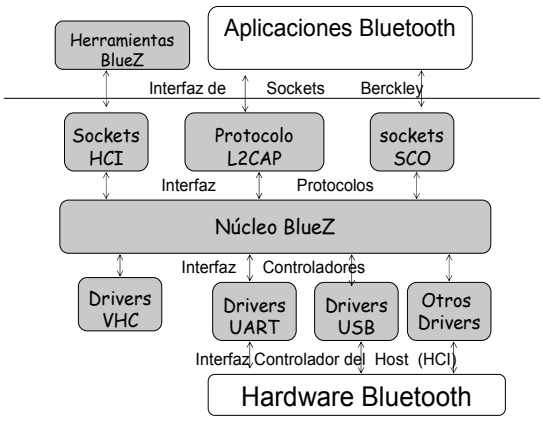


Figura 3-Pila de protocolos de BlueZ

4-Resultados

En este apartado se describen los resultados más relevantes de las pruebas.

Al evaluar los resultados obtenidos no tenemos que olvidar que el medio de transmisión utilizado por Bluetooth es un canal radio y como en todas las comunicaciones inalámbricas tendremos presente fenómenos de multitrayecto, directividad de las antenas e interferencias que afectarán a las medidas realizadas.

4.1- Tasa binaria a nivel de enlace

Como se comentó anteriormente el objetivo de esta prueba es determinar el ancho de banda útil máximo a nivel *L2CAP,* estudiando cómo influyen la distancia entre el dispositivo transmisor/receptor y el ambiente en el que se desarrollan las comunicaciones.

Primer Escenario

En la figura 4 están representados los resultados obtenidos en el primer escenario para los distintos tipos de paquetes. Este entorno es el más favorable para las comunicaciones inalámbricas. Las medidas realizadas en este escenario nos servirán como referencia, pudiendo determinar la degradación en las comunicaciones que provocan factores como la existencia de obstáculos o dispositivos interferentes que utilicen la misma banda de frecuencias.

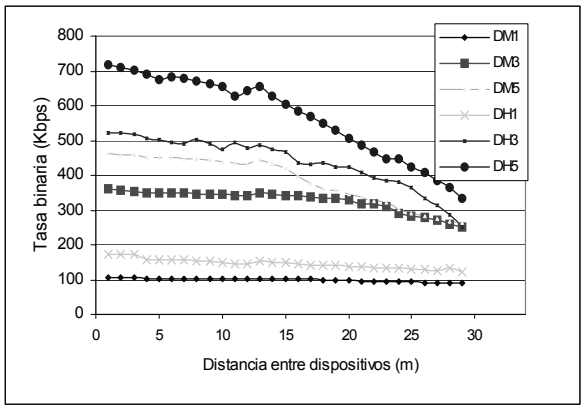


Figura 4-Tasa binaria a nivel de enlace en función de la distancia entre el dispositivos- Primer escenario

Si nos fijamos en la figura 4 observamos que para distancias cortas en todos los tipos de paquetes tenemos una tasa constante muy próxima a la indicada en el estándar, pero a partir de una distancia umbral la tasa binara comienza a descender de forma lineal.

La distancia umbral y la tasa de descenso de la curva dependerá del tipo de paquete que estemos utilizando. Si empleamos paquetes DM, que utilizan códigos de protección contra errores, las comunicaciones se verán menos afectadas por la distancia que si utilizamos paquetes DH, que al no poder corregir errores si tienen un error en destino se rechazan y se tienen que retransmitir. Podemos hacer un análisis similar si atendemos a la longitud del paquete. Los paquetes que ocupan 5 ranuras, aunque *a priori* tengan más capacidad para transmitir información también serán más susceptibles a errores, lo que provocará una retransmisión del paquete en los casos que no se puedan corregir los errores por el código de redundancia.

En la tabla 2 queda reflejada la anterior reflexión. En ella se indica el punto hasta donde se midió un ancho de banda constante, la tasa de descenso de la tasa binaria llegado ese punto y el área de cobertura para cada paquete, si entendemos como tal una degradación del 20% de la tasa inicial. Por ejemplo, para paquetes DM5 obtenemos un ancho de banda constante hasta los 13m, a partir de esta distancia la tasa binaria descenderá a un ritmo de 8,65 kbps/m, llegando a un descenso del 20% de la tasa inicial a 16m.

Segundo Escenario

Si comparamos los anteriores resultados con los obtenidos en el segundo entorno de pruebas, observaremos que aunque tenemos un comportamiento cualitativo similar al caso anterior: el área de cobertura y tasa de variación se mantienen prácticamente en el mismo valor; cuantitativamente la tasa binaria efectiva medida en este último escenario es mucho menor para todas las distancias, llegando a tener un descenso del 36% cuando empleamos paquetes DH5 (se paso de 717,396 kbps a 462,04 kbps). Los únicos paquetes que no se ven afectados son los DM1 y DH1, estos paquetes son más robustos que el resto por dos motivos: al ser más cortos las probabilidades de tener un bit erróneo o de que se produzca una interferencia con otro paquete son menores que para el resto.

En la figura 5 están representados los resultados de las medidas realizadas. Las curvas en este caso no son tan suaves, hay fluctuaciones y en muchos casos obtenemos una tasa binaria menor a distancias mas cercanas entre dispositivos que a distancias mayores. Esto se debe al efecto de las reflexiones y el multitrayecto.

Tipo de Paquete	Distancia Ancho Banda Constante (m)	Pendiente del tramo lineal (kbps/m)	Área Cobertura (Degradación 20%)
DM1	23 m	--	>29 m
DM3	20 m	--	25 m
DM5	13 m	-8,65	16 m
DH1	29 m	-11.15	21 m
DH3	15 m	-13.14	19 m
DH5	8 m	-16.48	15 m

Tabla 2-Resumen de los datos de cobertura y características de los resultados medidos en el primer apartado

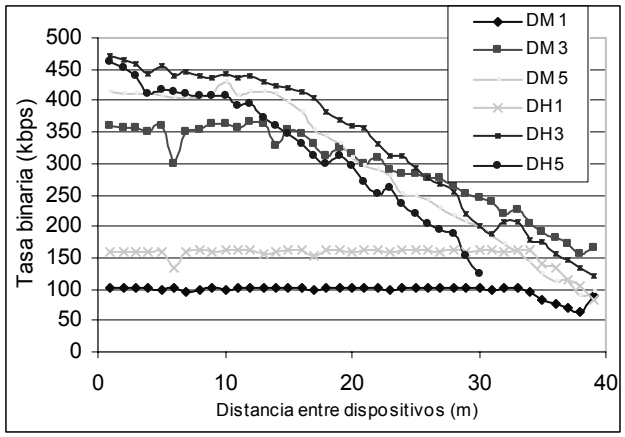


Figura 5- Tasa binaria a nivel de enlace en función de la distancia-Segundo escenario

Tercer Escenario

Mediante las pruebas en este escenario intentaremos evaluar el impacto que distintos tipos de obstáculos (muros, armarios, sillas) tienen sobre la tasa binaria efectiva. En la figura 6 se representan los resultados obtenidos al medir la tasa binaria efectiva para los distintos tipos de paquetes.

Estos resultados están caracterizados por las fluctuaciones debidas a los obstáculos que se encuentra la señal en su camino, de forma que nos encontramos con una tasa binaria más o menos uniforme dentro del recinto dónde se encuentra el nodo fijo. Pero cuando se traspasan las fronteras y cruzamos un muro de gran espesor (18m) obtenemos un descenso fuerte para todos los tipos de paquetes utilizados en las medidas. A raíz de los resultados podemos determinar que en escenarios de este tipo, la zona de cobertura vendrá determinada en gran medida por la topología del escenario dónde se desarrolle la comunicación, siendo un parámetro importante a tener en cuenta a la hora de diseñar un servicio que utilice una tecnología inalámbrica como es el caso de Bluetooth.

Al igual que en el caso anterior, las diferencias entre emplear paquetes de 3 y 5 slots son mínimas, por lo que es más eficiente utilizar paquetes más cortos, evitando la ocupación del medio de forma innecesaria.

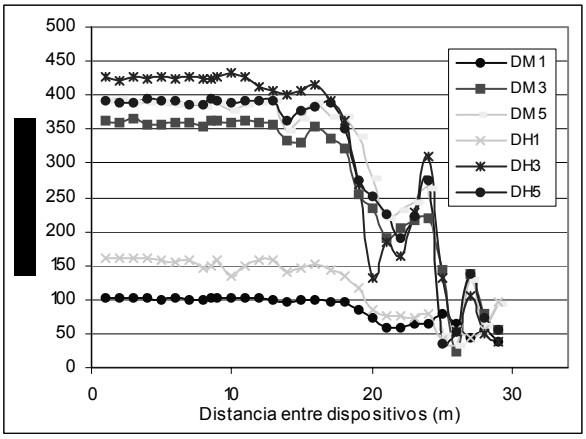


Figura 6-Tasa binaria a nivel de enlace en función de la distancia-Tercer Escenario

Cuarto escenario

Con las medidas realizadas en este escenario se pretende determinar el efecto que tendrían sobre la tasa binaria las comunicaciones de otra *piconet* activa en la misma área de cobertura que la utilizada para tomar las medidas. Para evaluar mejor el efecto de la interferencia se ha utilizado la configuración más desfavorable, colocando el transmisor de la *piconet* interferente al lado del dispositivo receptor bajo test.

Como comentamos en el segundo apartado, Bluetooth utiliza técnicas de salto de frecuencia para evitar interferencias. Por lo tanto, *a priori* dos *piconets* activas que compartan área de cobertura no deberían interferir una en la otra al utilizar distintos patrones de salto. Tan sólo se producirá interferencia cuando en un momento determinado la secuencia aleatoria utilizada por ambos dispositivos coincida en frecuencia, lo que no es muy probable, por ello se ha repetido la prueba 20 veces realizando una media. Los resultados se muestran en la gráfica 7 dónde está representada una comparación para cada tipo de paquete los resultados obtenidos con y sin interferencias de otra *piconet* Bluetooth.

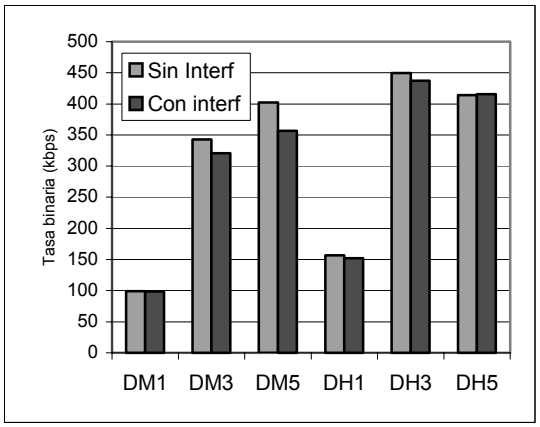


Figura 7-Comparación tasa binaria primer/cuarto escenario.

Conclusiones de las medidas realizadas

A través de los resultados anteriores hemos podido constatar que el escenario donde se vayan a desarrollar las comunicaciones es un factor crítico y es importante tenerlo en cuenta a la hora de desarrollar una aplicación con Bluetooth. Del entorno de comunicaciones dependerán en gran medida las capacidades en términos de ancho de banda, tanto en el valor máximo disponible para el transporte de datos, como en la estabilidad.

De forma que mientras en un espacio abierto sin interferencias no tendremos ningún problema al implantar un servicio que necesite un ancho de banda de 650 kbps, este servicio no será viable en un entorno cerrado.

Otro factor importante a la hora de diseñar un servicio con Bluetooth es la elección del tipo de paquete utilizado en la comunicación. Los resultados de las pruebas nos muestran que los paquetes más robustos al incremento de distancia y a entornos ruidosos son los DM al incorporar códigos de protección contra errores. El tamaño del paquete también que repercute en el resultado. Mientras que en espacios abiertos es efectivo utilizar paquetes que empleen 5 slots, para otro tipo de escenario deja de ser rentable provocando una ocupación innecesaria del espectro e interferencias.

En la tabla 3 hemos resumido los resultados obtenidos para paquetes que ocupan tres slots a una distancia de 7 m, incluyendo la desviación típica de la muestras tomadas en cada prueba.

Los resultados obtenidos en las pruebas son muy similares a los presentados por trabajos que han abordado el mismo tema, pero que han utilizando herramientas de simulación (*"Bluefog", "BlueHoc,"*). Aunque en nuestras pruebas se midieron valores de tasa binaria ligeramente menores que en los estudios realizados por Leopol y Fernandez en [1,2] en los que se utilizaban simulaciones, el comportamiento cualitativo y cobertura coinciden con estos estudios analíticos.

Entorno de Pruebas	DM3 (kbps)		DH3 (kbps)	
	$\overline{DM3}$	σ_{DM3}	$\overline{DH3}$	σ_{DH3}
1º Escenario	349,92	15,721	491,68	35,935
2º Escenario	350,64	34,56	446,88	45,55
3º Escenario	360,604	36,497	426,217	39,289

Tabla 3-Resumen valor medio / desviación para distintos escenarios.

4.2- Tasa binaria a nivel TCP

En este apartado mostraremos los resultados del ancho de banda disponible a nivel TCP en función de la distancia.

La figura 8 muestra los resultados obtenidos en el primer y segundo escenario, donde se pone de manifiesto la diferencia entre ambos escenarios. En un entorno cerrado hay un descenso de un 35% con respecto a uno abierto. Este resultado es paralelo al del apartado anterior, pero con una tasa de descenso mayor, además disminuye el rango de distancias con una tasa binaria constante debido al efecto del control de flujo del protocolo TCP.

Los mismos comentarios son aplicables al tercer escenario, la tasa binaria dentro del recinto donde se encuentra el equipo transmisor es aproximadamente constante con un valor de 360 kbps. Pero cuando salimos del recinto y hay por en medio obstáculos más contundentes, como pueden ser los muros, el ancho de banda desciende de forma abrupta con desconexiones constantes entre dispositivos cuando la distancia supera los 15 m. Los resultados de este escenario están representados en la figura 9.

Al repetir la prueba en el último escenario, como en el caso anterior, las medidas no se vieron afectadas por la actividad de otra *piconet* activa.

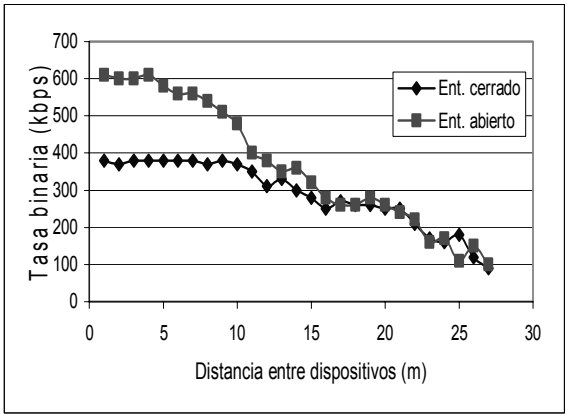


Figura 8- Tasa binaria TCP en función de la distancia primer /segundo escenario

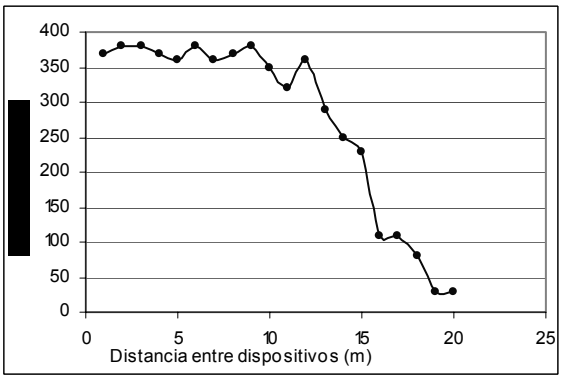


Figura 9- Tasa binaria TCP en función de la distancia tercer escenario

4.3-Tiempo medio de acceso

En la tabla 4 se recogen los resultados obtenidos al calcular el tiempo de acceso al medio en el segundo escenario para distancias de 1m, 10m, 15m, 20m y 29m. Si los presentamos en la siguiente tabla podemos observar en primer lugar que la distancia entre los dos nodos que se quieren comunicar casi no tienen influencia, puesto que el factor principal que afecta al tiempo de *Inquiry* es la correspondencia entre la frecuencia a la que se manda el mensaje de *Inquiry* y la frecuencia de escucha del otro dispositivo. También podemos observar un leve aumento del tiempo de *Inquiry* al realizarse la búsqueda en un entorno más ruidoso como es el segundo escenario de pruebas.

Otro cálculo interesante que podemos hacer es el de la probabilidad de tiempo de acceso acumulada para las medidas realizadas en entorno abierto con distancia entre los dos nodos de 1 m, representada en la figura 10. En ella podemos observar que para el 50% de los casos el tiempo de acceso es de 2 sg.

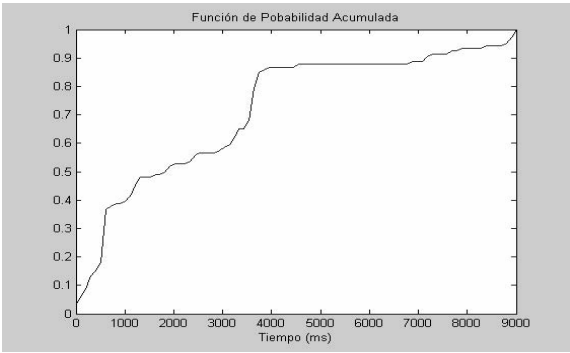


Figura 10- Función de probabilidad acumulada del tiempo de acceso en un entorno abierto

Distancia (m)	Entorno Abierto (Escenario 1)		Entorno Cerrado (Escenario 2)	
	$\overline{T_{inq}}$	σ_{Tinq}	$\overline{T_{inq}}$	σ_{inq}
1	2,37 sg	0,61	2,89 sg	1,52
10	2,13 sg	0, 75	2,81 sg	1,32
15	2,45 sg	1,3	2,63 sg	1,66
20	2, 68 sg	0,95	2,90 sg	1,84
29	2,52 sg	1,81	3,28 sg	1,92

Tabla 4- Tiempo medio de acceso

5-Conclusiones

En este artículo se han presentado y analizado los resultados de diversas pruebas realizadas con el fin de determinar las posibles capacidades de Bluetooth.

A través de nuestro contacto con dispositivos Bluetooth y plataformas de diseño software, hemos podido constatar que Bluetooth empieza a ser una tecnología madura, incorporada en una gran variedad de equipos que pueden interoperar entre sí. Esta penetración en el mercado trae consigo la aparición de nuevos servicios.

Uno de los logros conseguidos en la definición del estándar han sido la flexibilidad de su arquitectura, preparada tanto para adoptar protocolos y servicios ya definidos como para utilizar otros definidos por el

usuario, lo que ofrece grandes posibilidades al diseñador. Pero también hemos detectado algunas debilidades entre las que destacamos el ancho de banda disponible que en ocasiones es bastante reducido para los requisitos de las aplicaciones actuales.

En las pruebas realizadas se ha comprobado que el entorno en el que se desarrollan las comunicaciones, es un factor muy importante a tener en cuenta a la hora de dimensionar una red o diseñar un servicio que utilice esta plataforma. Las capacidades de Bluetooth dependen en gran medida del escenario de la comunicación y las posibles interferencias que existan en éste. El despliegue de aplicaciones en entornos en los que existan numerosos obstáculos es viable, manteniendo parte de su funcionalidad.

Debido a las limitaciones de cobertura, Bluetooth es una tecnología idónea para ser utilizada en aplicaciones dependientes de contexto, como puede ser guiado de museos, descubrimiento de servicios en dispositivos dependientes de contexto, postes de información turística, etc.

Se ha comprobado que la incorporación de protocolos de nivel superior, como pueden ser TCP/IP, es viable, facilitando la adopción en entornos tradicionales basados en arquitectura TCP/IP.

BlueZ es una herramienta muy potente que ofrece una gran cantidad de posibilidades a los desarrolladores. Gran parte su potencia se debe a su arquitectura modular, lo que permite acceder a distintos niveles de la torre de protocolos según la exigencia de la aplicación.

Una vez realizado el estudio, se ha planteado la ampliación del estudio añadiendo algoritmos de mejora para TCP o utilizando calidad de servicio, así como su posible utilización en ámbito docente.

Agradecimientos

Este trabajo ha sido parcialmente financiado por la Comisión Europea a través del proyecto Daidalos ((FP6-2002-IST-1-506997) y del proyecto SANTNEX (FP6-2002-IST-507052).

Referencias

[1] Martin Leopold,,"Evaluation of Bluetooth communication: Simulation and experiments". University of Copenhagen, 2002.

[2] D. Ferrandez Bell, J.C. Cano, P. Manzano, "Evaluating Bluetooth performance as the support for context-aware applications" Universidad Politectica de Valencia, 2003.

[3] M. Caldari, M. Conti, P. Crippa, G. Marozzi, "SystemC Modeling of a Bluetooth Transceiver: dynamic management of packet type in a noisy channel" Italy 2003.

[4] Urban Bilstrup Per-Arne Wiberg. "Bluetooth in Industrial Environment".School of Information Science, Computer and Electrical University.Halmstad, Septiembre 2000.

[5] N. Chevrollier y N.Golmie. "Techniques to Improve Bluetooth Performance in Interference Environments". National Institute of Standards and Technology Gaithersburg, Mariland 2003.

[6] G. Ennis. "Impact of Bluetooth on 802.11 Direct Sequence". IEEE 802.11 Working Group Contribution, IEEE, 802.11-98/319, Septiembre 1998.

[7] N. Golmie, R.E Van Dyck and A. Soltanian. "Interference of Bluetooth and IEEE 802.11: Simulation modeling and perfomance evaluation". Procceddings of de Fourth ACM International Workshop on de Modeling, Analysis, and Simulation of Wireless and Mobile Systems, Rome Italy (Julio 02).

[8] Bluetooth Special Interest Group Especificación del Sistema Bluetooth, Corev1.1.URL:http://www.bluetooth.com/pdf/ Bluetooth_11/SpecificationsBoock.pdf Febrero 2001.

[9] Personal Area Networking Profile. Bluetooth Special Interest Group Specification of the Bluetooth System. Diciembre 2002.

[10] Grupo de trabajo del IEEE 802.15 URL http://grouper.ieee.org/groups/802/15.

[11] N. Golmie. "MAC Performance Evaluation Process in a Coexistence Environment" IEEE 802.11 Working Group Contribution, IEEE 802.15-99/117, Septiembre 99.

[12] BlueZ – Pila de protocolos Bluetooth oficial de Linux. URL:http://bluez.sourceforge.net Enero 2005.

[13] E. Balatti, L Marzegalli, M. Vitiello. "Increasing TCP/IP Perfomance Over Home Wireless Networks". OPNETWORK 2001,

[14] Ling-Jyh Chen, Ruey-Lung Hsiao, "Implementación of TCP of Bluetooth". UCLA Computer Science Departamente, Los Angeles USA, 2001.

[15] L. Chen, R. Kapoor "Enhancing Bluetooth TCP Throughput via Link Layer Packet Adaptation", UCLA Computer Science Departamente, Los Angeles USA, 2004.

[16] A. Das. A. Ghose, A. Razdan, "Enhancing Perfomance of Asynchronous Data over the Bluetooth Wíreless Ad-Hoc Network", IBM Indian Research Laboratory, Indian Institute of Technology, Marzo 2001.

[17] Bluetooth Network Encapsulation Protocol Specification (BNEP). Bluetooth Special Interest Group, Diciembre 2002.

[18] Netperf. Evaluación de prestaciones en redes de comunicaciones. URL http://netperf.org. Enero 2005.

[19] AXIS OpenBT Stack, Pila de protocolos Bluetooth para Linux URL: http://sourceforge.net/projects/openbt/ visitado en Marzo 03.

[20] AFIXX Bluetooth Protocol Stack For Linux. URL http://affix.sourceforge.net, Dic 2003.

Arquitectura de gestión de un operador neutral Wi-Fi

Jaume Barceló, Anna Sfairopoulou, Jorge Infante, Miquel Oliver, Carlos Macián

Grupo de Investigación en Tecnologías y Estrategias de las Telecomunicaciones. Universidad Pompeu Fabra
Passeig de Circumval.lació, 8 .
08003 – Barcelona
Teléfono: 93 542 29 42 Fax: 93 542 25 17
E-mail: {jaume.barcelo,anna.sfairopoulou,jorge.infante,miquel.oliver,carlos.macian}@upf.edu

Abstract*. In this paper the management architecture for a Wireless Neutral Operator is analysed. We describe the basic principles and components of a Neutral Network (an access network infrastructure that belongs to different entities and is shared according to a memorandum of understanding signed by all the entities) and the modules needed for the system administration and support, such as service assurance, configuration and usage management. We also analyse how the different modules are structured. Experimental results have been obtained using a real system implementation in the Open Network of the University Pompeu Fabra.*

1 Introducción

El panorama de las redes de acceso ha cambiado notablemente con la proliferación de las redes Wi-Fi, en los más diversos ámbitos (en los hogares, aeropuertos, grandes superficies comerciales, etc.) Estas redes, aunque en principio capaces de dar conectividad para un amplio abanico de servicios telemáticos, se usan habitualmente para dar acceso a Internet. Históricamente, el acceso a la red, el transporte de los datos a través de ella e incluso los servicios a los que se accede (Internet, portales o cualquier otro) eran ofrecidos por el mismo operador de telecomunicaciones, que integraba toda la cadena de valor en una única empresa. El modelo de operador neutral que describe este artículo pretende separar el acceso a las redes de los servicios proporcionados por éstas, desagregando así la cadena de valor y permitiendo mayor flexibilidad, tanto tecnológica como comercial.

Sin embargo, este modelo crea una situación en la que ni el proveedor del servicio de valor añadido ni el gestor de la red de acceso Wi-Fi están en condiciones de gestionar la red de transporte, al contrario que en el modelo de integración vertical. Esto plantea una serie de problemáticas para la gestión de la red y del servicio, la descripción de cuya solución ocupa la parte central del artículo.

El resto del documento está estructurado como sigue: La presentación de los principios básicos del modelo, así como sus antecedentes y ventajas se describen en el segundo apartado del artículo.

El tercer apartado entra en los detalles de implementación del módulo de gestión del operador neutral, diferenciando cuatro aspectos: Las conexiones de los usuarios, gestión de fallos y prestaciones de la red, configuración de los distintos elementos que la integran y las medidas de uso y su aplicación para facturar a los distintos actores implicados.

El modelo aquí descrito ya ha sido implantado con éxito en un entorno ciudadano y universitario, proporcionando control de acceso, compartición de infraestructuras y niveles de servicio diferenciados a las entidades participantes. También está prevista la extensión del modelo a más universidades y la implantación de un operador neutral en otros entornos. Los detalles del despliegue se encuentran en el cuarto apartado del artículo. Finalmente, el capítulo 5 concluye el artículo.

2 El modelo de operador neutral Wi-Fi

El modelo de operador neutral Wi-Fi surgió ante la necesidad de compartir infraestructuras de acceso y sus costes de despliegue y mantenimiento [1]. Los modelos verticales tradicionales imponen que la misma entidad, normalmente un operador, despliegue una infraestructura de red de uso exclusivo y proporcione servicios finales sobre ella a los clientes del operador.

El objetivo del operador neutral es dotar de mayor flexibilidad al modelo y reforzar a los usuarios y a los proveedores de servicios. La idea es ofrecer una red de acceso *abierta*, a la que se puedan conectar todos aquellos usuarios que lo deseen y también todos los proveedores que dispongan de servicios para ofrecer a los usuarios finales.

2.1 Antecedentes

El germen del presente modelo se encuentra en la experiencia desarrollada por KTH en Estocolmo bajo la denominación Stockhölm Open [2]. En dicho proyecto se ha desplegado una infraestructura de red troncal neutral a la cual se conectan redes locales tanto de fibra óptica y cable como inalámbricas para suministrar servicios de acceso a Internet.

Originalmente Stockhölm Open estaba orientada a incorporar redes de infraestructura fija desplegadas en los domicilios, para lo cual contaba con la colaboración de Stockab, consorcio público de gestión de alquileres de propiedad inmobiliaria de titularidad pública. Actualmente el modelo de Stocklhöm Open se ha extendido a redes públicas inalámbricas que cubren ubicaciones en las dependencias de KTH, Kista, algunas ubicaciones en la ciudad y el archipiélago de Estocolmo [3]. Un total de cuatro proveedores de servicio Internet (ISPs) suministran servicio a través de Stockhölm Open.

La experiencia de Stockhölm Open se ha tomado como referencia inicial para el operador neutral desarrollado en la Universitat Pompeu Fabra, partiéndose del software de libre disposición y de los modelos de negocio definidos en Estocolmo para adaptarlos a las necesidades propias.

2.2 Principios básicos

En el modelo propuesto aparecen cuatro actores principales: los usuarios, las islas Wi-Fi, el operador neutral y los proveedores de servicios (Fig. 1).

Los usuarios son los clientes de los servicios de telecomunicación. Provistos de un terminal, se conectan a la red de acceso y a través de ella alcanzan a los proveedores de servicio.

Las islas Wi-Fi consisten en zonas de cobertura inalámbrica a partir de puntos de acceso conectados a la red troncal del operador neutral. El operador neutral es la parte central del modelo y actúa de ente gestor entre usuarios, islas Wi-Fi y proveedores. Es su responsabilidad el despliegue y gestión de la red troncal, y supervisar que el resto de actores cumple los acuerdos.

Los proveedores son aquellas empresas o entidades que ofrecen servicios de valor añadido a los usuarios. Ejemplos de estos servicios son acceso a Internet, e-government, voz sobre IP inalámbrica, etc.

2.3 Ventajas del operador neutral

La aplicación del modelo de operador neutral respecto a un despliegue monolítico de red Wi-Fi y servicio de acceso a Internet tiene ventajas que se pueden desglosar en las siguientes para cada colectivo:

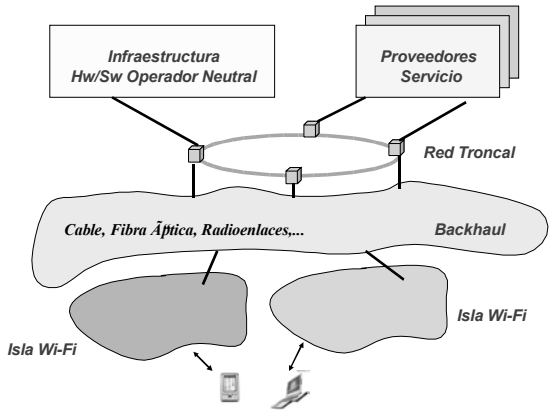


Figura 1: Arquitectura general de operador neutral

Ventajas para los usuarios finales

La existencia de un operador neutral permite que el usuario pueda:

- Acceder a servicios de diversos proveedores a través de una única red definida por la cobertura del operador neutral, seleccionando el proveedor de servicio en función de las características y precio de éste.
- Acceder puntualmente (usuario ocasional), si su proveedor de servicio se encuentra conectado al operador neutral, desde un punto distinto al habitual sin tener que contratar específicamente acceso Wi-Fi para esta ubicación, ya que su contrato con el proveedor de servicio cubrirá todas las ubicaciones donde exista cobertura del operador neutral.
- Acceder a servicios gratuitos que consistirán en consultas a webs institucionales (por ejemplo, la guía de Barcelona de www.bcn.es) o de asociaciones ciudadanas o empresas que deseen suministrar estos servicios de forma gratuita a los posibles usuarios.

Ventajas para los proveedores de servicio

La oferta del operador neutral para los proveedores de servicio se concreta en:

- Comercializar servicios de telecomunicaciones pudiendo hacer uso de la conectividad con alta capilaridad que ofrecen los servicios de operador neutral, siendo posible de esta manera centrar el núcleo del negocio y las propuestas innovadoras en la oferta de servicio.
- Disponer de una base de posibles clientes conformada por todos los usuarios ubicados en las localizaciones cubiertas por los operadores de red acogidos al modelo de operador neutral.
- Conectar a la red del operador neutral y ofrecer sus servicios manteniendo su autonomía, es decir teniendo la posibilidad de seguir comercializando servicios sobre una infraestructura de red propia o ajena y ofrecer estos mismos servicios a través de la red gestionada o acogida el operador neutral sin tener que asumir restricciones o limitaciones en el resto del negocio.

Ventajas para los proveedores de red

El modelo de operador neutral resulta de especial interés para los operadores cuyo núcleo del negocio se apoya en la disponibilidad de infraestructura de red en ubicaciones privilegiadas con un elevado número de usuarios potenciales con perfil de uso intensivo de servicios de telecomunicaciones. Para este tipo de operadores, la conexión al operador neutral plantea la ventaja de permitir rentabilizar la red desplegada a través del uso que hagan de ésta los proveedores que

comercialicen sus servicios a través de la infraestructura de red acogida al operador neutral.

De esta manera, los operadores de red pueden centrar sus actividades en el despliegue y mantenimiento de la infraestructura de red, disponiendo de una base de clientes sin por ello tener que abordar la provisión y mantenimiento de servicios finales, ni tener que realizar campañas de marketing y gestión de canales de comercialización de las cuales se responsabilizan los proveedores de servicio. Este modelo resulta especialmente ventajoso para pequeños operadores de red que no disponen de capacidad financiera para abordar la comercialización de servicios en un mercado en competencia, pero que sí disponen de un conjunto, que no necesariamente tiene que ser extenso, de ubicaciones privilegiadas en las que se prevea un elevado número de usuarios potenciales.

Como ejemplos típicos de este tipo de actores en el contexto de redes inalámbricas, se tendrían aquellas empresas que alcanzan acuerdos para el despliegue y explotación de redes inalámbricas en estaciones, aeropuertos, hoteles, o centros comerciales y que desean alcanzar el número máximo posible de usuarios de la red.

Al igual que en el caso de los operadores de servicio para la gestión de clientes, los operadores de red mantienen su autonomía en la gestión de su propia infraestructura y en los posibles usos alternativos o simultáneos que deseen dar a la red, tanto comerciales como de uso privado, siempre que respeten los acuerdos de calidad de servicio que se establezcan con el operador neutral.

Ventajas para las administraciones publicas

La gran mayoría de las administraciones públicas disponen de proyectos e iniciativas para acercar la sociedad de la información a los ciudadanos, estando centradas gran parte de ellas en potenciar los servicios de e-government a través de Internet.

El modelo de operador neutral da la oportunidad a las administraciones públicas de ofrecer este tipo de servicios a los ciudadanos en las ubicaciones en que estén presentes los operadores de red acogidos al modelo de operador neutral. Bajo el esquema de operador neutral la administración pública actuaría como un proveedor que se conecta al operador neutral y ofrece a los ciudadanos estos servicios (en general de forma abierta, pero la administración, si así lo desea, puede establecer modelos de pago).

3 Arquitectura de Gestión del operador neutral.

En el operador neutral desarrollado en el marco del proyecto COSF (Catalunya Oberta Sense Fils) [4], la gestión del operador neutral se realiza a cuatro niveles.

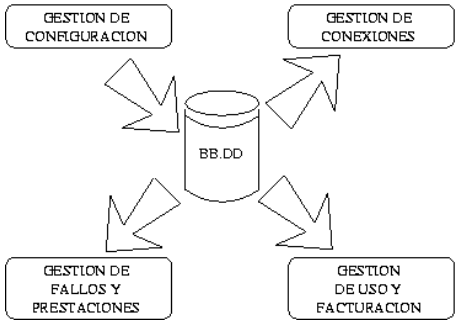


Figura 2: Módulos de gestión de un operador neutral

Por un lado la gestión de conexiones permite que un usuario, tras un proceso de registro y autenticación, acceda a los servicios de uno de los proveedores. La gestión de fallos y prestaciones mantiene información crítica del estado de los elementos básicos de la red, así como de los servicios que presta ésta. Con la gestión de uso, se guardan estadísticas y se crean informes de tráfico y de uso de cada operador e isla Wi-Fi, necesarias para la facturación. Finalmente, con la gestión de configuración se puede mantener una base de información actualizada sobre los equipos, entidades de red, servicios, islas, proveedores de servicio, etc., que se utiliza en el resto de módulos para realizar el resto de actividades indicadas teniendo en cuenta la configuración de los equipos, redes y servicios que forman el operador neutral (Fig. 2).

3.1 Arquitectura de Gestión de Conexiones.

Crear una red inalámbrica es fácil: basta conectar un punto de acceso (que proporciona la interfaz radio) a una red clásica y ya se pude disfrutar de la tecnología sin hilos. Sin embargo los problemas llegan pronto, especialmente en el caso de las organizaciones complejas: aparecen amenazas a la seguridad en la transmisión de la información y a la estabilidad de la red tradicional.

En la mayoría de casos, la arquitectura más adecuada consiste en crear una red fija virtual dedicada exclusivamente a los puntos de acceso para conexión inalámbrica, también denominada red de acceso. De este modo aislamos los posibles problemas de seguridad que puedan surgir. Si se dispone de una infraestructura de red adecuada, esta solución presenta múltiples ventajas adicionales. Entre ellas, cabe destacar la posibilidad de utilización de los denominados 'servidores de acceso' (*access server*) para la conexión de la red inalámbrica al resto de nuestra red. A diferencia de los *routers* y *firewalls* tradicionales, un servidor de acceso valida los usuarios (por ejemplo con una clave de paso y contraseña almacenadas en un servidor LDAP, o bien RADIUS) como paso previo antes de permitir el tráfico de datos a través suyo.

Una arquitectura basada en una red y un servidor de acceso soluciona la mayor parte de los problemas que la tecnología Wi-Fi presenta a los administradores de red. Sin embargo, si se pretende compartir la red de acceso se requieren soluciones que den respuesta a retos adicionales. La validación de usuarios y contraseñas en un servidor LDAP centralizado y único puede resultar insuficiente: cada proveedor tendrá sus propios usuarios, sus propios métodos de identificación y sus propios servicios. Una solución basada en una arquitectura de red de acceso abierta permite dar respuesta a estas necesidades (ver Fig. 3).

El modelo de operador neutral desarrollado consiste en una red de nivel 2 a la cual se conectan servidores y puntos de acceso. Cada proveedor de servicios utiliza un esquema de direccionamiento de nivel 3 (IP) distinto dentro de una misma red de nivel 2, más una red IP *default* utilizada por los usuarios no registrados.

El elemento clave de la red es el servidor de configuración [5], que almacena la asociación entre terminales y proveedores de servicio. Asimismo es necesario un servidor de acceso para cada proveedor de servicio para controlar las conexiones de sus usuarios.

La configuración de los equipos terminales se realiza mediante el protocolo *dhcp* , por tanto el usuario que desee conectarse al operador neutral deberá configurar su interfaz Wi-Fi para que utilice dicho protocolo. Se precisa un servidor *dhcp* externo a la red, al actuar el servidor de configuración como *dhcp relay* para el funcionamiento del sistema.

Al arrancar un equipo terminal, o cuando éste detecta una red Wi-Fi se realiza una petición *dhcp* por parte del equipo terminal que se desea conectar. El *dhcp relay* del servidor de configuración recoge dicha petición y la reenvía al *dhcp server*. En caso de que el terminal todavía no esté asociado a ningún proveedor (por ejemplo, un terminal que se conecta por primera vez al operador neutral) es asignado a la red *default*.

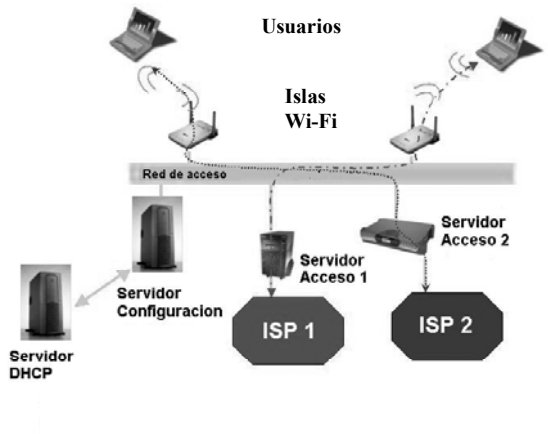


Figura 3: Gestión de conexiones

La red *default* tiene la característica de tener el servidor de configuración como *default gateway* y un *lease-time* de renovación de los parámetros IP muy bajo (algunos segundos). Dicho servidor permite únicamente el tráfico DNS y redirige todo el tráfico por los puertos *http* y *https* hacia una página de selección de proveedor existente en su propio servidor web.

Cuando el usuario abre un navegador y solicita una URL cualquiera, se produce la resolución DNS y a continuación la petición *http*. El servidor de configuración intercepta dicha petición y, en respuesta, presenta al usuario un formulario en el cual pide al usuario que seleccione uno de los proveedores de servicio disponibles. Una vez el usuario escoge un proveedor se debe esperar a la expiración de la dirección IP (en nuestro caso unos pocos segundos). Al ser renovada ésta se cambia por otra del proveedor elegido, así que el usuario es redirigido hacia el servidor de acceso del proveedor que, a su vez, será el *default gateway* de la red IP de dicho proveedor dentro la red de acceso .

La elección de proveedor de servicio realizada por el usuario se guarda en la base de datos del servidor de configuración, de tal modo que cuando el terminal realiza una nueva petición de IP se le asigna automáticamente el proveedor de servicio preseleccionado.

El *access-serve*r permite el tráfico únicamente de usuarios autenticados y autorizados. Típicamente, cuando un usuario intenta navegar, se le presenta un formulario en el que se le solicita un identificador y contraseña. Éstos se verifican utilizando un fichero local, o un servidor de autenticación como LDAP, RADIUS o KERBEROS, a elección del proveedor de servicio. Si el usuario tiene permiso, se le abre el *firewall* del *access server*. Un *access server* debe monitorizar si el usuario permanece activo para cerrar el *firewall* una vez concluida la sesión. Existen diversas implementaciones de *access server* [6][7], unas propietarias y otras de código libre, cada una con sus ventajas e inconvenientes.. Varias de ellas pueden coexistir en una misma red neutral. En UPF se ha optado por software de código abierto tanto para el servidor de configuración como para el servidor de acceso.

3.2 Gestión de Fallos y Prestaciones.

En una red neutral, donde participan diferentes y heterogéneos actores, es muy importante disponer de un módulo de gestión para supervisar el funcionamiento correcto de todos los elementos involucrados, independientemente de quién sea su propietario. Con este objetivo hemos creado el módulo de Gestión de Fallos y Prestaciones, un módulo que usando el protocolo SNMP (Simple Network Management Protocol), nos permite realizar un control remoto de los elementos y puntos de interconexión del operador neutral y obtener información en tiempo real de su configuración . De este modo el administrador de la red puede tener una imagen completa del funcionamiento de cada componente (puntos de acceso, switches, etc.),

detectar problemas y usar esta información para crear estadísticas e informes de calidad de servicio.

El módulo de Gestión de Fallos y Prestaciones interroga periódicamente a todos los componentes principales de la red y guarda la información obtenida en una base de datos para procesar off-line. Los datos extraídos de las MIBs de cada equipo (MIB2, ver tabla 1) son los que describen el estado del componente (activo o no), el tiempo durante el cual ha estado funcionando, el porcentaje de paquetes erróneos por paquetes enviados, el número de usuarios conectados, etc. Con esta información, el sistema nos permite gestionar los fallos y las prestaciones y detectar incumplimientos de los contratos de Nivel de Servicio (SLAs) de los diferentes actores.

Se ha desarrollado una herramienta gráfica que permite la monitorización en tiempo real del sistema, mostrando alertas en caso de que un componente presente un funcionamiento erróneo.

Tabla 1: MIBs usados por los modulos de Gestión de Fallos y Gestión de Uso

GESTION DE FALLOS		
	MIB_NAME	**Comentario**
Access Point	IfOutUcastPkts	Información para cada una de las interfaces de punto de acceso (wireless, Ethernet etc) + información general de disponibilad (upTime) del AP
	IfOutDiscards	
	IfInUcastPkts	
	IfInDiscards	
	IfAdminStatus	
	IfOperStatus	
	sysUpTime	
Switch	IfAdminStatus	Información general de disponibilidad del switch (upTime), paquetes enviados, paquetes descartados i estadode cada interfaz en el que están conectados islas Wi-Fi o ISPs
	IfOperStatus	
	ifLastChange	
	IfOutUcastPkts	
	IfOutDiscarded	
	ifInUcastPkts	
	ifOutDiscarded	
	SysUpTime	
GESTION DE USO		
AP	UserTable	Numero de usuarios conectados en cada instante a cada AP
Switch	IfOutUcastPkts	Información de trafico para cada interfaz del Switch donde se conecta una isla Wi-Fi o un ISP
	IfOutDiscarded	
	ifInUcastPkts	
	ifOutDiscarded	

También se pueden generar informes de errores y historiales del estado del sistema, que pueden ayudar a identificar el orígen de problemas experimentados por los usuarios.

3.3 Gestión de Configuración.

Dada la coexistencia de diversos proveedores de servicios en la misma arquitectura, será necesario gestionar las altas y bajas y disponer de herramientas que comprueben si el proveedor cumple los requisitos para conectarse al operador neutral.

La incorporación progresiva de nuevas islas Wi-Fi que proporcionan nuevas zonas de cobertura, obligan a disponer de un sistema común para gestionar el crecimiento. Tiene que permitir la alta de nuevas islas, comprobar que cumplen los requerimientos exigidos y validar su funcionamiento.

Se dispone de una base de datos donde se mantiene toda la información necesaria sobre los proveedores de servicio y las islas Wi-Fi, como por ejemplo los datos de contacto para cada proveedor de servicio, el número de puntos de acceso que pertenecen a cada isla, los puntos de conexión de una isla/operador con la red neutral (los puertos del switch que les están asignados) etc . Estos datos son utilizados por el resto de módulos de gestión.

3.4 Gestión de Uso y facturación

También es necesaria una herramienta para gestionar el uso de la red de acceso. Se mide el número de usuarios registrados por cada proveedor, el tráfico generado por dichos usuarios, así como el tráfico generado por cada una de las islas Wi-Fi y como éste se reparte entre los diferentes proveedores de servicio (ISPs). El objetivo es obtener información sobre la utilización que hacen los usuarios y evitar usos fraudulentos o no permitidos para proponer un sistema que permita efectuar una facturación por uso de los recursos y elementos del sistema.

El módulo de Gestion de Uso tiene como base el mismo código que el módulo de Gestion de Fallos, ya que su funcionamiento es muy parecido. Usando como herramienta principal el protocolo SNMP, obtenemos toda la información necesaria interrogando los switches y puntos de acceso de la red de manera periódica. Los MIBs usados para la gestión de uso se pueden consultar en la tabla 1. Cada puerto de un switch está asociado a un proveedor o a una isla Wi-Fi, así que recogiendo información sobre el tráfico observado en este puerto podemos tener información sobre el tráfico del proveedor/isla Wi-Fi correspondiente. La información se almacena en una base de datos MySQL y se procesa usando la herramienta gráfica mencionado anteriormente para crear todos los informes y visualizar historial y estadísticas de uso. Esta información se combina luego con información almacenada en la base de datos específica para la facturación (tarifas, requerimentos mínimos de uso y disponibilidad etc) y la información de Gestor de Configuración y se usa para facturar a los proveedores de servicio y pagar a las islas WiFi.

4 Experiencia de implantación del operador neutral

El modelo del operador neutral ha sido implantado con éxito en la Universidad Pompeu Fabra y ya ha sido extendido a entidades próximas como el IDEC (Institut d'Educació Contínua) con el objetivo de compartir áreas de cobertura. Además se contempla en un futuro inmediato su extensión al resto de universidades catalanas dentro del proyecto Catalunya Oberta Sense Fils financiado por la Generalitat de Cataluña y realizado por la fundación I2CAT, de la cual la Universitat Pompeu Fabra forma parte.

Otro proyecto asociado consiste en desplegar un operador neutral en el distrito tecnológico 22@ de la ciudad de Barcelona con la participación y colaboración de empresas públicas y privadas [9].

4.1 Implantación del Sistema en la Universitat Pompeu Fabra

UPF creó su red de acceso inalámbrica el año 2001, mediante la instalación de 10 puntos de acceso que ofrecían conectividad Wi-Fi desde las bibliotecas de la universidad. Esta red ha ido creciendo paulatinamente hasta los 35 puntos de acceso actuales, ampliándose la cobertura a auditorios, salas de reuniones y otros espacios de interés.

El mecanismo de autenticación establecido inicialmente, basado en el registro manual de las direcciones MAC de las tarjetas Wi-Fi de los usuarios en un servidor *dhcp*, ofrecía serios problemas tanto de escalabilidad como de seguridad. Por este motivo a finales del año 2003 se inició el estudio de la mejora del mismo [8]. Se buscaba una solución basada en código abierto que permitiera compartir la red de acceso con otras organizaciones próximas a la universidad y el uso de la misma por parte tanto de usuarios como de proyectos de investigación en el ámbito de las tecnologías inalámbricas.

La solución implantada ha permitido ofrecer a casi dos millares de usuarios registrados una red de acceso de gran cobertura y facilidad de uso al tiempo que garantiza la seguridad del resto de redes, que se encuentran protegidas por los *access servers*. Además permite un trato diferenciado de los distintos tipos de usuario usando distintos proveedores.

La red neutral de la UPF permite la conexión a Internet a través de RedIris a los usuarios de la universidad y la conexión a la red experimental i2cat a los investigadores que lo precisan.

También se ha instalado un proveedor para proporcionar acceso a Internet a los participantes a congresos que usan las instalaciones de la universidad. Este proveedor se activa de manera puntual durante los días que dura el congreso.

Tabla 2: Número de MACs registradas en cada proveedor

PROVEEDOR	*# MACs*
UPF	1538
i2CAT	6
IDEC	104
CONGRESS	204
Total:	1852

Además la red de acceso está compartida con otra institución en la cual se imparten los masters de postgrado de la UPF (IDEC, Institut d'Educació Contínua) beneficiándose ambas de una extensión de las zonas de cobertura. IDEC dispone de su propio servidor de acceso y gestiona sus propios usuarios, así como los servicios que les ofrece.

La tabla 2 muestra el número de direcciones MAC registradas con cada proveedor en Abril de 2005. Este número se corresponde aproximadamente con el número de usuarios que han utilizado el servicio. El ritmo de crecimiento observado en los últimos meses es de unas 150 MAC/mes.

4.2 Interconexión con otras Universidades catalanas.

En el marco del proyecto COSF (*Catalunya Oberta Sense Fils*) se ha desarrollado un kit que contiene el software y la documentación necesaria para el despliegue de un operador neutral con cobertura más amplia que la propia UPF.

Este kit se ha distribuido a cinco universidades seleccionadas, para su estudio e incorporación progresiva al operador neutral ya existente. Una vez concluida la integración, los estudiantes y personal de cada universidad podrá acceder a las redes de su propia universidad desde las instalaciones de cualquiera de las otras.

4.3 Otras implantaciones y proyectos futuros.

El Ayuntamiento de Barcelona en colaboración con la Universitat Pompeu Fabra ha lanzado la iniciativa 22@WiFi [9] cuyo objetivo principal es desplegar en el distrito 22@ un operador neutral Wi-Fi para la prestación de servicios de telecomunicaciones. Mediante esta iniciativa se pretende:

- Enriquecer la oferta de servicios de telecomunicación en el distrito 22@.
- Facilitar la entrada en el mercado de operadores de servicios, reduciendo las barreras de entrada al mercado.

- Realizar una experiencia piloto de compartición de infraestructuras de red entre operadores de servicio.
- Poner a disposición de los ciudadanos y de los operadores de servicios acceso inalámbrico en ubicaciones públicas gestionadas por el ayuntamiento de forma transparente, abierta a todos los proveedores de servicio y no discriminatoria.

Para llevar a cabo el proyecto, el Ayuntamiento promocionará la implantación de un operador neutral Wi-Fi que utilizará como red troncal la red de fibra óptica municipal disponible en este distrito, desplegará puntos de acceso públicos y ofrecerá la gestión del tráfico entre los puntos de acceso y los proveedores de servicio que se acojan a esta iniciativa para facilitar el establecimiento de un modelo de operador neutral abierto a la incorporación del mayor número posible de proveedores de servicio e islas Wi-Fi.

5 Conclusiones

El modelo del operador neutral es una idea innovadora que proporciona interesantes beneficios a todos los actores participantes, derivados principalmente de las ventajas de compartir infraestructuras aportadas por diversos pequeños operadores de red entre operadores de servicios gracias a la intermediación del operador neutral. En este artículo hemos presentado sus antecedentes y ventajas, y hemos dado los detalles de una posible implementación de operador neutral.

El operador neutral ya ha sido implantado con éxito en el ámbito universitario y ha dado sus primeros frutos en términos de control de acceso y compartición de infraestructura y áreas de cobertura de entidades distintas. En estos momentos se está extendiendo el operador neutral existente al resto de universidades catalanas.

También se está trabajando en la mejora del modelo y en la adaptación del mismo a otros entornos, en los que tengan cabida otros actores como ayuntamientos o empresas privadas.

Para ello se ha desglosado la gestión del operador neutral en cuatro módulos. El primero proporciona las herramientas básicas para permitir la interconexión entre usuarios y servicios. El segundo ofrece gestión de fallos y prestaciones, permitiendo el acuerdo de niveles de servicio y su posterior verificación. La gestión de uso analiza la participación de los distintos actores, principalmente en términos de tráfico generado o recogido, abriendo las puertas a políticas de facturación. Finalmente, el módulo de gestión de configuración contiene información sobre los elementos de la red, redes, servicios y actores implicados, necesaria para sustentar el resto de módulos.

Agradecimientos

Los trabajos descritos en este artículo se han podido realizar gracias al soporte de la Generalitat de Catalunya y la fundación I2CAT, que han financiado el proyecto Catalunya Oberta Sense Fils, marco en el cual se han realizado y están extendiéndose los módulos que componen el operador neutral, así como el ayuntamiento de Barcelona, que han financiado el análisis de su aplicación en el distrito 22@ de Barcelona y también a través del proyecto SIMA (TIC2003-07279-C02-01). Los autores desean agradecer también las correcciones y sugerencias de los revisores.

Referencias

[1] R. Battiti, R. Lo Cigno, F. Orava, B. Pehrson. "Global Growth of Open Access Networks: from WarChalking and Connection Sharing to Sustainable Business" WMASH'03, Septiembre 2003, San Diego, USA.

[2] F. Orava, M. Wall, D. Liberal, O. Lundström, T Rautiainen, P. Samlin, M. Setterberg. "SwedenOpen.net Final Report". Communications System Design 2003.Mayo, 2003. KTH, Suecia.

[3] The Stockholm Open.net Project. http://www.stockholmopen.net

[4] M. Oliver. "COSF: Descripció de l'Arquitectura d'Operador Neutral". Proyecto COSF 2004.

[5] A. Escudero, B. Pehrson, J.-O. Vatn E. Pelleta, P. Wiatr. "Wireless Access in the Kista-IT University" .11th IEEE Workshop on Local and Metropolitan Area Networks (LANMAN2001), Marzo 2001, Boulder, USA.

[6] M.Hedenfalk."Access Control in an Operator Neutral Public Access Networks". Proyecto Final de Carrera, Royal Institute of Technology (KTH), 2002.

[7] NoCatAuth. Proyecto NoCatNet. http://www.nocat.net

[8] J. Barcelo. "Implantació d'una Arquitectura Oberta d'Interconnexió de Xarxes Sense Fils". Proyecto Final de Carrera. Universidad Pompeu Fabra (UPF).

[9] "22@WiFi:Descripción y Plan de Proyecto". Julio, 2004.

Computer Support in Group-based Learning. A Meta-model Contribution to Educational Modelling Languages

Manuel Caeiro Rodríguez, Martín Llamas Nistal, Luis Anido Rifón
Departamento de Ingeniería Telemática. Universidad de Vigo
ETSI de Telecomunicación. C/ Maxwell S/N. Campus Universitario.
36310 – Vigo (Pontevedra)
Teléfono: 986 814 173 Fax: 986 812 116
E-mail: {Manuel.Caeiro; Martin.Llamas; Luis.Anido}@det.uvigo.es

***Abstract**. The paper describes the basis of a meta-model to support the design of computer-supported Group Based Learning (GBL) processes and scenarios. This proposal is conceived as an initial step to create a new Educational Modelling Language (EML). An EML provides a framework of elements that supports the description of any design of a teaching-learning experience in a formal way, involving learning participants, resources, tasks, scenarios, etc. EMLs are related with CSCL scripts, which are proposed as tasks processes that aim to facilitate GBL. The goal of these initiatives is to produce formal pedagogical descriptions that may be processed by appropriate software engines to provide automatic management and coordination of the elements.*

1 Introduction

During the last years, educational and technological developments have promoted the appearance of a many *Group-based Learning* (GBL) systems. GBL involves learners working together in a series of activities in order to complete a joint learning task. Recent pedagogical theories and research findings consider that learning can be more effective when: (i) students are actively involved in the process of knowledge creation [1]; (ii) learning takes place in a realistic context [2], and (iii) learners work together and help each other to understand what they are learning [3]. These assumptions have led to a shift away from traditional lecturing towards GBL formats in many settings [4]. Alongside these developments in pedagogical paradigms, technical advances have enabled many initiatives aiming to enhance the education in group-based settings through the use of computer support [5].

Therefore, the design of GBL systems supported by computer technologies involves the management of pedagogical and technical issues. This paper is focused on the technological support that can be provided by computer technologies to enable the design of GBL settings. We analyze the elements involved in the design of GBL [6] and consider the interactions that need to be supported in group-based systems [7]. Then, we propose a meta-model devoted to describe GBL settings, which support the elements and requirements analyzed. This meta-model is the first stage to develop a new *Educational Modelling Language* (EML) [8] that enables the truly design of collaborative learning scenarios.

The rest of the paper is organized as follows. In the next section the GBL design is considered by identifying the main involved elements. In section 3, we consider the different kinds of interaction involved in GBL settings. Next, in section 4, a meta-model is proposed to enable the design of GBL. In section 5 we use the meta-model to describe a design solution. Next, this proposal is situated in the context of EMLs. The paper ends with some conclusions.

2 Group-based Learning Design

Currently, the design of Computer-Supported GBL (CSGBL) is often based on subjective decisions regarding tasks, pedagogy and technology [6]. Classic instructional design focuses on individual learning outcomes and tries to control instructional variables to create a learning setting that supports the acquisition of a specified knowledge or skill. But in computer-supported GBL the use of groups complicate this view. There is a proposal of five critical elements that affect interaction in GBL [6]:

- *Learning Objectives*. The learning objectives affect the expected interaction to be maintained in a learning environment. It is possible to differentiate in a continuum ranging from "*closed skills*", relatively fixed skills that can be learned separately (e.g. a procedure for long divisions), to "*open skills*" (e.g. argumentation, negotiation). These different kinds of skills require different kinds of interactions. A *closed skill* mainly requires interaction consisting in reactive remarks. In contrast, *open skills* require different kinds of interaction, not only reactive.
- *Task Type*. The task type also can be depicted on a continuum ranging from "*well-structured tasks*", which often require the application of a limited number of rules or principles and have one correct solution, to "*ill-structured tasks*", which have a considerable degree of uncertainty regarding the rules and principles that can be applied and often have no clear solution.

- *Level of Pre-structuring*. This addresses the observation that collaboration sometimes develops spontaneously, but more often it does not. Therefore, a continuum is proposed that addresses the level to which interaction is pre-structured in advance by either a teacher or designer ranging from "*high pre-structuring*", to "*low pre-structuring*". The right level of pre-structure is a key point in computer-supported GBL design: too much structure may result in forced artificial interaction, but no structure may result in fragmented inefficient interaction.
- *Group Size*. Group size has influence in the kind of interactions that appear in collaboration. Work in small groups will produce much more interactions and will require less coordination than in large groups. Furthermore, a larger group requires more effort from group members to achieve common ground.
- *Computer Support*. It is possible to distinguish between two approaches to support GBL activities: *structuring* and *regulating* collaboration [9, 10]. In face-to-face situations, teachers *structure* and *regulate* student interaction by, first, preparing the lesson plans and setting up the group work, and then, intervening in the collaboration when they feel it is necessary. In computing contexts, *structuring* approaches aim to create favourable conditions for learning by designing and scripting the situation *before* the interaction begins. They attempt to define the structure of the learning experience by varying the characteristics of the participants (e.g. the size and composition of the group), the availability and characteristics of tools and communication media, and the nature of the task (e.g. writing, problem solving). *Regulation* approaches support collaboration by taking actions *after* the interaction has begun. They compare the dynamically changing state of student interaction to a model of "desired" interaction, and intervene when discrepancies between these two states are discovered.

3 Computer-support in Group-based Interaction

GBL involves the interaction among a group of participants that, in our context, has to be supported by a computational system. To work in group, people need to exchange information (i.e. to *communicate*), to act together in a shared space (i.e. to *co-operate*), and to organize themselves (i.e. to *coordinate*). Computer technologies provide different solutions that support these interactions. In this section we consider the different interaction types together with the tools and technologies that can be used to support group-based interaction.

As we have introduced, we identify three ways of interaction in group-based systems, which we named perspectives: (i) *communication*, groups functionality related to unconstrained and explicit communication among the participants; (ii) *co-operation*, groups functionalities related to storage and access to share data; and (iii) *coordination*, related to the management of dependencies among participants and activities. This set of perspectives is proposed accordingly to groupware studies [7, 11, 12], coordination theory principles [13, 14], Human-Computer Interaction models [15], and e-learning proposals [16]. The classification is not complete, in the sense that it is possible to find more interaction ways that do not fall in any of these perspectives. For example: (iv) *awareness*, concerned with participants gathering information from others in the space and form their activities; and (v) *group decision making*, referred to the way in which a collaborative decision is adopted. But, we focus the attention on the first three ones, due to space constraints.

3.1 Communication

Communication encompasses the process of transfer and exchange of information that takes place between participants [7, 16]. Communication is a basic perspective involved in any collaborative situation. Typical communication tools are: e-mail, desktop conferencing systems, chat, whiteboard, etc. The typical functionalities of these tools are: sending and receiving a message, joining and leaving a conference, managing mailing lists, etc.

These tools usually consider several mechanisms to control the access and utilization of their operations by participants. *Conference* and *conversational* models are the underlying control mechanisms:

- The *conference model* describes whether only two or more people can communicate and how that communication is initiated, and if more than two party conferences are allowed, how new people join the conversation, whether it is possible within a muti-party conversation to talk privately to some subset of the group, and so on. The conference model must also specify whether all participants can transmit/receive, and if it is not possible how one switches from transmitting to receiving.
- The *conversation model* describes what are the conversational moves allowed in the communication, how participants take turns in performing these conversational moves, what are appropriated conversational replies to the moves, how the groupware can help the user manage each conversation, and manage multiple conversations. There is a broad rangeCSGBL developments that consider tools with specific *conversation models* (e.g. CSCL scripts [6, 10]).

In *synchronous communication* tools the emphasis is on the *conference model*. It is assumed that the participants themselves will manage the conversation (i.e. participants have to interpret the conversational moves and follow the appropriate group protocols for such moves). In *asynchronous communication* the

emphasis is on the *conversation model*. Because there may be a long period between one conversational message and its reply, the system may provide help to its users (e.g. providing the context of a received message indicating the preceding ones, listing all messages that need reply, stating that some types of messages need no reply but should be processed automatically, etc.)

3.2 Co-operation

Sometimes the collaboration among a group of people is centred on the access and change of a shared set of data [7, 16]. In these situations, the goal of the collaboration is the construction of this shared data, namely the *artefact*. In co-operation[1], we group the set of interactions related to the storage and manipulation of artefacts. Examples of systems that provide these functionalities are shared editors, virtual whiteboards, shared repositories, etc. The functionalities involved are:

- *Control access rights to the objects*. Not all participants have the same rights to the objects that make up the *artefact* or the same rights to perform some operations onto these objects.
- *Control of simultaneous access to the artefact*. Some tools allow for simultaneous changes to the *artefact*. This poses the problem of maintaining the consistency of the *artefact* (e.g. if two simultaneous and contradictory changes are submitted, how will the tool perform them?).
- *Versioning of the artefact*. In some applications it is important to store stable situations of the *artefact* during the process and to allow the *artefact* to be restored to such stable situations.
- *Storage of time stamp and author information on objects of the artefact*. Some tools allow a user to view just the changes since the last logging.
- *Floor control*. Some tools use a mechanism of floor control to avoid simultaneous access to the *artefact*, for example a classroom blackboard. At each time only one user has the right to change the *artefact* (the participant that has control of the floor). Other users may request the floor which will be granted by the system as soon as the participant that has the floor relinquishes it.

It is interesting to note that some of these functionalities may be related with the models introduced in the *communication* perspective. For example, the "*floor control*" functionality is very similar to the "*conversational model*", and the first two *co-operation* functionalities are related with the "*conference model*". This allowed us to consider the same kind of solution to support the representation of both behaviours.

3.3 Coordination

Coordination is considered as the process of managing dependencies among tasks or activities[2] [7, 13]. In general, coordination also involves the management of dependencies among resources (e.g. artefacts, tools), but in this category we exclusively gather functionalities related to activity coordination. The resource dependencies were already considered in the *communication* and *co-operation* perspectives (e.g. the *conference* and *conversation* models).

In groupware, the necessity of activity coordination mechanisms has been the centre of a heated discussion [14, 16]. At one side, there are normative models that try to *structure* the collaboration by restricting the interaction between participants and their activities. This limits the flexibility of collaborative systems. At the opposite side, there are those advocating that collaborative systems should take flexibility to the extreme, leaving the coordination burden to the users and simple *mediating* in the interaction. This augments the coordination workload to users. In spite of this discussion, there is a trend to conciliate both ideas, arguing that both kinds of activities are "seamlessly meshed and blended in the course of real world" [17].

In CSGBL, there exists a similar problem, related with the first three design criteria described in section 2 (*learning objectives*, *Task Type*, and *level of pre-structuring*). There is a large group of collaborative activities (mainly *closed skills*, *well structured task*s, and *high level of pre-structuring*) whose tasks depend on each other to start, to be performed, and to end. On the other hand, coordination does not need to appear explicitly in some kind of activities (*open skills*, *ill structured task*s, and *low level of pre-structuring*) such as those realized by means of chats or audio and videoconferences.

The basic functionalities of coordination are centred on the execution (or enactment) of a plan, or a sequence of tasks (e.g. a process). The coordination is devoted to assure that an instance of the plan follows the predefined specification. Such functionalities related to enactment are:

- *Enabling an activity* once its preceding activities have terminated.
- *Notification to the users* that they may start a particular activity or that it is late.
- *Inspecting the current stage of a process*. Some systems allow privileged users to obtain information about the process state (e.g. which

[1] Note that we use the form "co-operation" to distinguish it from "cooperation". In groupware, "cooperation" is usually considered as a kind of collaboration where a global goal is split in several sub-goals, and participants are assigned to different sub-goals.

[2] The terms activity and task are used interchangeably. Alain Wisner makes the important distinction between 'task' and 'activity' [18]: "*Tasks are what managers set - they are the prescribed work. Activity is what people actually do*"

tasks have been completed, and when, and by who, and which activities are being carried on).

- *Dynamic alteration of a process description to cope with surprises*. Changes to the plan are important in dealing with unexpected situations, which were not taken into consideration when the plan was conceived.
- *Helping participants to manage their work*. A support system may help the actor by displaying the list of activities already performed, displaying the deadlines, and allowing the user to choose which task to do.

Another important group of functionalities centres on defining the *process* accordingly to a *coordination model*. The *coordination model* has two components: a component that deals with the *modelling* of the process and one that deals with the *enactment*. The process is a predefined specification on how an endeavour will or should proceed. The process specifies the tasks and their goals, the participants who have to perform them, the objects and operations available in each task, the order in which the tasks should be performed, when the tasks should end, etc. The *coordination model* has to specify which of these components are fixed and which can be set by the user. The main concept of the *coordination model* is that of a *task*. Other important concepts are *role* and *actor*. A *task* is a potential set of *operations* (and the corresponding *objects*) that an *actor* playing a particular *role* can perform, with an intended goal.

4 A Proposal to Model the Design of Group-based Learning

The purpose of this section is to present a solution devoted to support the design of GBL settings. In the previous sections we described both the design issues and the interactions that should be considered in computer-supported GBL. The proposed solution consists in a meta-model that enables the description of GBL scenarios. It is based on workflow management, groupware, and human-computer interaction models and proposals. Especially, we follow key ideas from [19, 20, 21].

Currently, we are developing an associated graphical representation and an XML binding to support the composition of documents accordingly to the meta-model.

4.1 The Meta-model

The proposed meta-model tries to support the design issues and interaction perspectives introduced in the previous sections. Basically, it can be considered as a *coordination model* that enables the definition of communication and co-operation functionalities (e.g. *conference* and *conversation* models). The real *communication* and *co-operation* functionalities (e.g. send a message, co-write a document) are not considered in the meta-model specifically, but they are included through the inclusion of tools and services that provide the appropriate functionality.

The most basic concept of the meta-model, c.f. fig. 1, is the *Educational Scenario* (ES). An ES is intended to perform a certain task: *well-structured*, *ill-structured*, etc. (e.g. solve an exercise, read a document, support a learner working with a simulator), focused towards some educational objective: *open skill*, *closed skill*, etc. (e.g. to get some expertise, to get some knowledge) and considering certain restrictions (e.g. time conditions).

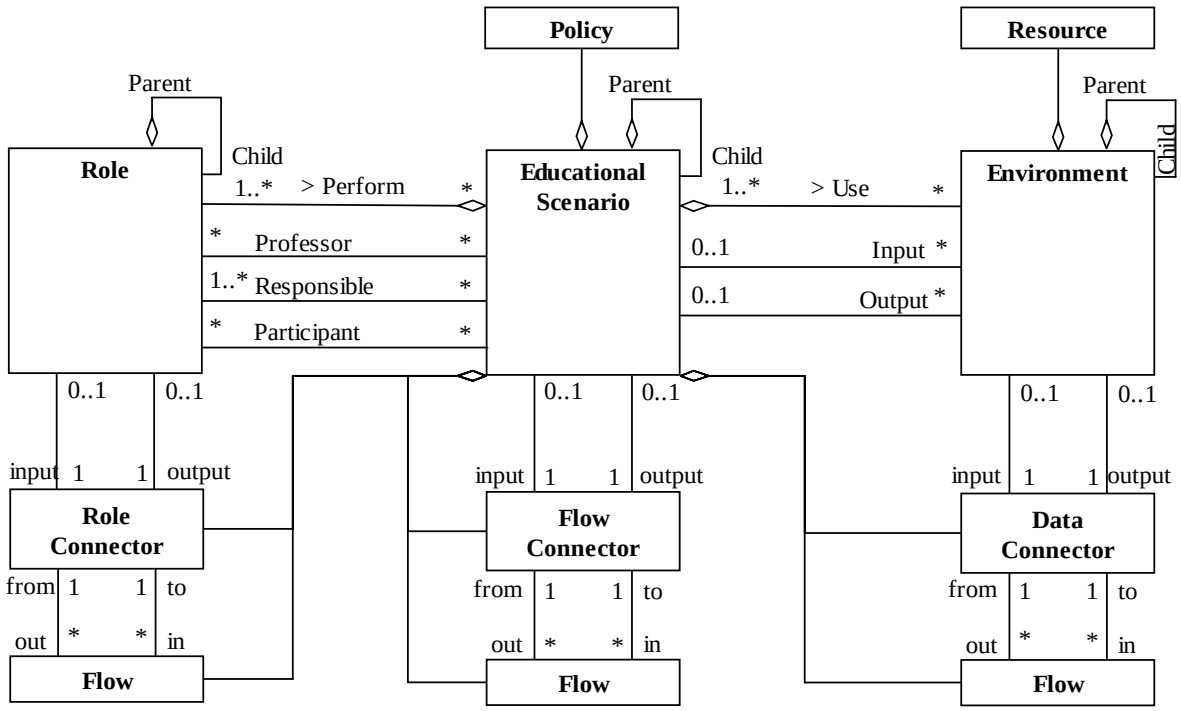


Figure 1: The proposed meta-model

The ES element is the aggregation point where all other elements are anchored. Furthermore, each ES constitutes a context of elements not accessible from other ESs. In this way it is possible to construct well-structured and separated work spaces. An ES aggregates the following elements:

- *Educational Scenarios*. An ES may contain other ESs devoted to breakdown a general goal in several sub-goals. All the ESs present the same organization. Relationships (*Connectors* and *Flows*) are considered to relate the elements of a Parent ES with the elements of a Child ES and the elements of Child ESs among them.
- *Roles*. The ES needs to consider some participants, at least one, to *Perform* the intended tasks. The number of participants is not restricted, and participants can be arranged in group hierarchies as desired. We have considered three special roles: (i) the *Professor* role is not mandatory, but if it is included the role can modify the prescribed ES as desired and control its enactment; (ii) the *Responsible* role is mandatory (but may be performed by several participants), and it has authorization to define new child ESs, providing a breakdown of the current one; (iii) finally, there can be any number of *Participants*. In this way we support the description of groups of different sizes. Furthermore, we support the specification of group structures.
- *Environments*. The ES can consider some *Environments* containing the *Resources* to be *Used* to achieve the intended goal. *Environments* can be grouped hierarchical, but finally they will be composed by information documents (e.g. HTML pages, texts, figures, etc.) and tools (e.g. communication facilities, simulators, text processors, etc.). There may be two special kinds of *Environments* associated to an ES, namely *Input* and *Output* parameters. These parameters work in association with *Input* and *Output* conditions to decide when a certain ES need to be enacted of finished.
- *Connectors*. We consider three different kinds of connectors: *Role Connector*, *Flow Connector*, and *Data Connector*. Each one of them provides several mechanisms to relate the corresponding elements in ESs: (i) the *Flow Connector* is used to relate the sequencing among ESs (e.g. sequence, OR-join, AND-split, parameter conditions, temporal conditions, etc.); (ii) the *Role Connector* enables the transfer of participants through roles, supporting that a certain participant can perform different roles in different ESs (e.g. free, *Responsible* decides); (iii) the *Data Connector* supports the data flow among *Environments* and ESs (e.g. copy, transfer, synchronize, etc.). The *Connectors* use *Rules* to provide conditioned behaviors.
- *Flows*. They are simple links used to connect the different *Connectors*. *Flows* are associated with the *Connectors*, never with other elements (*Roles*, ES, *and Environment*).
- *Policies*. They are used to specify authorizations for the roles involved in an ES. *Policies* are used to define the permissions (*rights*, *obligations*, *prohibitions*, *dispensations*) that each role is provided to use operations in the elements and resources of the ES [22].

An ES is the main element of the meta-model, representing a unit of education at any level of granularity or specificity. Therefore, ESs play a central role, being both the concept for process structuring, and the corner stone on which reuse is promoted. We have taken this idea from the workflow domain, where it has been established that a system that does not include separate classes for atomic and composite tasks can more easily accommodate design-time and run-time evolution, making the language simpler and more efficient [20]. The approach is a mixture of block-based and flow-based workflow model. ESs are grouped via a hierarchy of ESs. Composite ESs are either *collections*, which have an unordered list of sub-items, or *networks* where some sub-items are interconnected. *Collections* and *networks* thus refer to models with different degree of structure. The interconnection network is made up of two types of elements: *Flow Connectors* and *Flows*. Each ES may have two *Flow Connectors* to describe the conditions to initiate and finish. This approach enables to construct designs with different *learning objectives*, *task types*, and *level of pre-structuring*. The *learning objective* is related with the goal of the ES. The *Task Type* depends on the nature of the task. The *level of pre-structuring* is based in the aggregation of ESs and the description of connections. Furthermore, we note that the structure of an ES may be modified during enactment.

The functionalities of the *communication* and *co-operation* perspectives would be supported by including in the environments appropriate *resources* (e.g. tools). In addition, it would be necessary to consider methods to configure and control the interactions among *resources* and *roles* and *resources* among them. *Resource* to *role* interaction is supported through policies. *Resource* to ESs will be supported by modelling the resource as a finite-state-machine. *Resource* to *resource* interactions are described by data connections.

5 An example

In this section we depict a small modelling example using some of the elements of the introduced meta-model. The example is based on the subject "*Ingeniería del Software I*", which is offered in the second year of the telecommunication engineering studies of the University of Vigo. In this subject, after some initial lectures, learners are proposed several practices that they have to carry out in pairs. In fig. 2 we present a model to articulate each one of such

practices, as an ES named *Software Engineering Practice*. Currently, we are developing the notation used in the figure. This notation is based in a workflow proposal [20]. In the figure, we have only included the definition of *Roles*, *Resources*, *sub-ESs*, and sub-ESs connections (*flow connectors* are represented by circles). Roles connections, data connections and *Policies* are not included to simplify.

The ES *actors* are defined in the upper part of the figure. As you can see, the *Software Engineering Practice* involves two main kinds of actors: *Pair* and *Tutor*. A *pair* is a group, made up of two actors: *Designer* and *Programmer*. These actors are transferred from the parent ES roles (i.e.: the subject actors: *Learners* and *Tutors*) using *Role Connectors* and *Flows*. We have decided to assign *Parent Learners* to *Designers* and *Programmers*, and *Parent Tutors* to *Tutors* directly. In the subject model, we have considered a sequence of three *Software Engineering Practice* ESs. The *Pair* participants are maintained in each of the three *Practice* ESs, but changing the *Learners* role. Namely, the *Designer* in the first *Practice* will be the *Programmer* in the second, and the *Programmer* in the first will be the *Designer* in the second. The role of each *Learner* in the third *Practice* is decided according to the previous results. In relation with the children ESs, roles are transferred to each sub-ES directly, accordingly to the roles defined in it.

Sub-ESs are represented in the central part of fig. 2. In the example, the *Software Engineering Practice* ES is decomposed in five sub-ESs connected by *Flow Connectors* and *Flows* (they should be further specified to explicitly represent the connection conditions: time limits, Join conditions, etc.). The first sub-ES is devoted to design a solution to the proposed practice. This task has to be done by the *Designer* role of the *Pair*, using a *Design Environment* to produce a *Design* document. When the first sub-ES is finished, two sub-ESs are initiated: *Programming* and *Design Evaluation*. The *Design Evaluation* sub-ES is assigned to the *Tutor*, who has to review the proposed design. As result a *Design Evaluation* document is provided. This review may be used by the *Designer* to modify and update the initial *Design* document in the next sub-ES: *Design Revision*. In parallel with these two sub-ESs, the *Pair* is devoted to program the intended design in the *Programming* sub-ES. This task may be assigned to the *Programmer* only, but we prefer that both participants collaborate in the programming due to educational reasons. Furthermore, the *Pair* may decide to breakdown this task in several sub-tasks to organize and manage their programming activities (e.g. each member of the *pair* may program different functions and then they integrate it). This is enabled by allowing the *Programmer* role to specify new sub-ESs. When the *Pair* finishes the *Programming* ES, the *Program Evaluation* sub-ES is activated. This sub-ES is assigned to the *Tutor* to review and evaluate the final *Pair* products *Design* and *Program* documents to produce a *Program Evaluation*.

Environments are depicted in the bottom part of the figure. We use different notations to represent tools (e.g. *Design Environment*, *Programming Environment*), documents (e.g. *Practice Specification*, *Design*, *Program*, *Design Evaluation*, and *Program Evaluation*), and sub environments (*Test Files*). Documents have to be transferred between the different sub-ESs. We have not represented this transfer, but it is obvious considering the resources defined in each sub-ES. The *Practice Specification* document should be included in all the sub-ESs, but it is also not included due to space limitations. The only interesting point involves the *Programming* and *Design Revision* sub-ESs. Both of these ESs use the document *Design*, that may be modified and updated in the *Design Revision* sub-ES. Therefore, both documents must be synchronized when the *Designer* perform some change in the *Design Revision* ES. This connection should be represented in the figure using an appropriate *Data Connector* (*synchronize*).

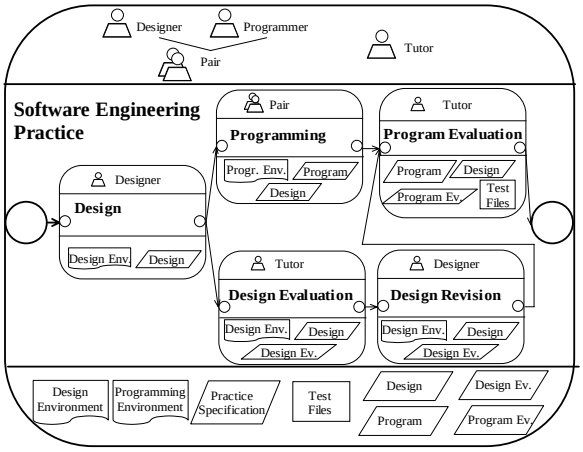


Figure 2: An example of a composed ES

6 Related Works

Our proposal is directly related to *Educational Modeling Languages* (EMLs) [8]. These languages were proposed some years ago to enable the design of educational designs. Accordingly to the CEN/ISSS WS-LT [23] survey of Educational Modeling Languages *"An EML is a semantic information model and binding, describing the content and process within a 'unit of learning' from a pedagogical perspective in order to support reuse and interoperability'*. Therefore, the purpose of EMLs, of which *IMS Learning Design* (LD) is the most outstanding proposal [24], is to support the description of diverse teaching-learning experiences (i.e. learning designs), embodying different kind of tasks, processes, persons, tools, and contexts.

The foundation of any EML is the meta-model that describes its elements and the relationships. In this way, the presented meta-model is the initial stage to

define our own EML or extensions to the LD proposal. Currently, the LD proposal is a meta-language that allows to codify *units-of-study* (e.g. courses, course components, programs of study), associating each element of content (e.g. texts, tasks, tests, assignments) with information describing its instructional strategy (e.g., roles, relations, interactions, and activities of students and teachers). LD is mainly concerned with supporting the coordination issues that take place in education: '*how to control the order of specific activities to be performed by humans and applications, and the use of resources*'. But, we consider that the IMS-LD language does not provide a good support to group-based issues, such as the presented in this paper: communication and co-operation perspectives, different levels of pre-structuring, group management, etc.

During the last years the people of the Open University of the Netherlands together with researches of other institutions have been promoting the development of tools and applications supporting the LD specification. The first developed tool was LAMS [25] in Australia. It is not compatible with LD but it is inspired in it. LAMS system enables the creation and running of certain educational processes in the form of sequences of learning activities. An architecture for a distributed LD authoring and runtime environment has been proposed by the so called Valkenburg group. RELOAD [26] (a funded project in the UK) has developed a LD editor and player. Its focus is on general tools that enable the fulfilment of all the LD elements, without the establishment of any predefined pedagogical approach. Also as part of the Valkenburg initiative the OUNL developed CopperCore [27] and CopperAuthor [28] to provide execution and authoring services, respectively. These services provide the main functionalities required for the management of LD models. They are proposed to enable that different institutions may develop particular end-user tools and applications satisfying their own requirements, but using the facilities provided by the Copper suite. In Canada, a research group at the University of Quebec has developed general graphical editor MOT+ [29] together with a runtime engine called Explor@-2 that is the basis for a LD runtime player. MOT+ has been specialized with added graphic objects to cover the IMS LD components. ASK Learning Designer Toolkit (LDT) [30] developed at the University of Piraeus (Greece) is a graphical tool supporting the authoring of learning activities that enables the development of complex LD scenarios.

In group-based settings, EMLs are related with CSCL Scripts [10]: "*A script is a story or scenario that the students and tutors have to play as actors play a movie script*". In group-based educational scenarios, these scripts are proposed as activity programs that aim to facilitate collaborative learning by specifying activities in collaborative settings, eventually sequencing these activities and assigning them to learners. But, currently there is no language or EML-related proposal that completely satisfies the design requirements involved in CSCL Scripts. As a consequence, our purpose is to contribute to the development of EMLs proposing a meta-model that supports such issues.

7 Conclusions

The EMLs approach to describe and support the design of diverse learning scenarios is going to play a key role in the future of e-learning systems. Our final goal is to contribute to the development of EMLs, focusing on GBL. In the paper we consider the requirements that must be considered in the design of computer-supported GBL settings. It is important to have in mind that GBL pedagogies involve certain particular conditions, which must be taken into account when computer-support is planned. One of the most important issues is the variability in the collaborative situations. GBL design may range along several criteria (mainly learning objectives, task type, and level of pre-structuring) that must be supported in any grade in order to provide a comprehensible support.

We have proposed a meta-model that enables the description of educational designs accordingly to the different possibilities. Furthermore, we have considered the possibility to modify educational designs during their execution or enactment. The meta-model core advantages are: (i) to provide a clear and simple way to articulate work; (ii) to enable different forms of structure; (iii) to support the interaction and coordination among participants, resources, and tasks; and (iv) to facilitate the flexibility of the final designs.

Acknowledgements

We thank Spanish "*Ministerio de Education y Ciencia*" for its partial support under grant "*MetaLearn: methodologies, architectures and languages for E-learning adaptive services*" (TIN2004-08367-C02-01).

References

[1] D. H. Jonassen, K. L. Peck, B. G. Wilson. *Learning with Technology: A Constructivist Perspective*. Upper Saddle River, NJ: Merril.

[2] H. S. Barrows, R. M. Tamblyn. *Problem-based Learning*. New York, Springer.

[3] L. S. Vygotsky. *Mind in Society: The Development of Higher Psychological Processes*, Cambridge, Harvard University Press (1977).

[4] J. van der Veen. *Telematic Support for Group-based Learning*. Ph. D. Thesis, University of Twente, Twente University Press (2001).

[5] D. H. Jonassen, S. M. Land. *Theoretical Foundations of Learning Environments*. New York, Lawrence Erlbaum Associates.

[6] J. W. Strijbos, R. L. Martens, W. M. G. Jochems "*Designing for Interaction: Six Steps to Designing Computer-supported Group-based Learning*" Computers & Education, 42, pp. 403-424 (2004).

[7] C. Ellis, J. Wainer "Groupware and Computer Supported Cooperative Work" In G. Weiss (Ed.) *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*, The MIT Press, pp. 425-457 (1999).

[8] R. Koper, *Modeling units of study from a pedagogical perspective – The pedagogical metamodel behind EML*, Open University of the Netherlands (2001).

[9] P. Jermann, A. Soller, A. Lesgold "Computer Software Support for Collaborative Learning" In J. W. Strijbos, P. Kirschner, and R. Martens (Eds.), *What We Know About CSCL in Higher Education*, Kluwer, pp. 141-166 (2004).

[10] P. Dillenbourg "Over-scripting CSCL: The Risks of Blending Collaborative Learning with Instructional Design" In P. Kirschner, (Ed.) *Three Worlds of CSCL: Can We Support CSCL?*, Kluwer, pp. 61-91 (2003).

[11] J. Lonchamp, "Process Model Patterns for Collaborative Work" Proceedings of the 15th IFIP World Computer Congress, Telecooperation Conference, Austria, 1998.

[12] A. B. Raposo, H. Fuks "Defining Task Interdependencies and Coordination Mechanisms for Collaborative Systems" Frontiers in Artificial Intelligence and Applications, pp. 88-103, 74, IOS Press (2002).

[13] T. W. Malone, K. Crowston "What is Coordination Theory and How Can It Help Design Cooperative Work Systems?" Proceedings of CSCW, pp. 357-370, 1990.

[14] K. Schmidt, C. Simone "Coordination mechanisms: Towards a conceptual foundation of CSCW systems design". CSCW, pp. 155-200, 4(2-3), (2000).

[15] D. Pinelle, C. Gutwin, S. Greenberg "Task Analysis for Groupware Usability Evaluation: Modeling Shared-Workspace Tasks with the Mechanics of Collaboration" ACM Transactions on Computer-Human Interaction, 10(4), pp. 281-311 (2003).

[16] A. B. Raposo, M. G. Pimentel, M. A. Gerosa, H. Fuks, C. J. P. Lucena "Prescribing E-learning Activities Using Workflow Technologies" CSAC'04, pp. 71-80, 2004.

[17] K. Schmidt, C. Simone "Mind the Gap! Towards a Unified View of CSCW" 4th Conf. on the Design of Cooperative Systems, 2000.

[18] A. Wisner "Situated cognition and action: implications for ergonomic work analysis and anthropotechnology" In *Ergonomics*, 38 (8), pp. 1542-57 (1995).

[19] N. Rusell, A. H. M. ter Hofstede, D. Edmond, and W. M. P. van der Aalst, *Workflow Data Patterns*, Technical Report, FIT-TR-2004-01, Queensland University of Technology (2004).

[20] H. D. Jørgensen, *Interactive Process Models* Ph. D. Thesis, Norwegian UST, (2004).

[21] W. M. P. van der Aalst, A. H. M. Hofstede "YAWL: Yet Another Workflow Language" Technical Report FIT-TR-2003-04, (2003).

[22] L. Kagal, *Rei: A Policy Language for the Me-Centric Project*. HP Labs Technical Report, HPL-2002-270 (2002).

[23] A. Rawlings, P. van Rosmalen, R. Koper, M. Rodríguez-Artacho, P. Lefrere, *Survery of Educational Modelling Languages (EMLs)*, Version 1, CEN/ISSS WS-LT (2002).

[24] R. Koper, B. Olivier, and T. Anderson, (Eds.) *IMS Learning Design Information Model*, IMS Global Learning Consortium (2003).

[25] J. Dalziel, "Implementing Learning Design: The Learning Activity Management System (LAMS)" ASCILITE, 2003.

[26] RELOAD, *Reusable eLearning Object Authoring and Delivery project website*, http://www.reload.ac.uk/ldesign.html, 2005.

[27] Coppercore project website, http://www.coppercore.org, 2005.

[28] CopperAuthor project website, http://www.copperauthor.org, 2005.

[29] MOT website, http://www.licef.teluq.uquebec.ca/francais/real/demot.htm, 2005.

[30] P. Karampiperis, D. Sampson, "A Flexible Authoring Tool Supporting Adaptive Learning Activities". CELDA 2004, Portugal, 2004.

Uso de técnicas de virtualización para mejorar la docencia en laboratorios de redes de comunicaciones

David Fernández Cambronero, F. Javier Ruiz Piñar, Fermín Galán Márquez,
Vicente Burillo Martínez, Tomás de Miguel Moro
Departamento de Ingeniería Telemática (DIT). Universidad Politécnica de Madrid (UPM)
ETSI de Telecomunicación. Ciudad Universitaria s/n 28040 MADRID
Teléfono: 915495700 Fax: 913367333
E-mail: david@dit.upm.es

Abstract. *This paper discusses about the use of virtualisation techniques in computer network laboratories, with the aim of reducing the deployment and maintenance costs, as well as to improve students' learning by allowing them to work on realer networking scenarios. Networking laboratories purely based on physical infrastructure are of great educational value but they involve a high configuration, operation and maintenance effort, as well as high equipment costs. With the help of virtualisation techniques, these overheads can be reduced without diminishing the educational value of the laboratory experiments. The paper presents a general introduction to virtualization, focused on UML (User Mode Linux) virtualization back-end and the VNUML (Virtual User Mode Linux) front-end tool developed at DIT-UPM. Besides, DIT-UPM computer network laboratory is presented, showing several different VNUML based virtualization scenarios, some of them already experimented.*

1 Introducción

En el ámbito de la enseñanza de la Ingeniería Telemática los laboratorios de redes tienen un papel muy importante, al permitir a los alumnos poner en práctica los conocimientos adquiridos en las asignaturas de teoría. En general, las prácticas pueden plantearse sobre escenarios de simulación y sobre escenarios basados en equipamiento real. Estos últimos tienen una gran importancia docente, puesto que permiten que los alumnos se familiaricen con la configuración y uso de equipos reales, realizando experimentos en los que observan la operación real de los servicios y protocolos de comunicación que han estudiado, a la vez que ejecutan pruebas de diagnóstico y prestaciones. Así, el valor del laboratorio de redes se ve grandemente reforzado si se dispone de una infraestructura de comunicaciones que permita construir escenarios experimentales variados, versátiles y de cierta amplitud.

Sin embargo, este planteamiento presenta ciertos inconvenientes. Por un lado, si se dispone de una infraestructura física que ofrece muchas posibilidades a la hora de plantear experimentación sobre ella, también se multiplica el esfuerzo de configuración, operación y mantenimiento que tal entorno demanda. Por otro lado, puede que la dotación concreta de los laboratorios no permita construir escenarios de prácticas adecuados a los objetivos docentes que se plantean, por limitación del número de recursos o del uso que se pueda hacer de ellos. En la actualidad, la potencia del hardware disponible (equipos cada vez más rápidos en términos de CPU, memoria, disco, etc) ha motivado el interés por las técnicas de virtualización y sus aplicaciones, con el objetivo principal de reducir costes de despliegue y gestión de lo que sería un sistema equivalente realizado de forma convencional con equipos reales. Los laboratorios docentes no son una excepción y la virtualización aporta grandes ventajas a su implementación, como se describirá a continuación.

El artículo se organiza como sigue. En la sección 2 se expone la situación de los laboratorios de redes en la ETSI de Telecomunicación de la Universidad Politécnica de Madrid (UPM) para ilustrar los problemas esbozados antes sobre un caso concreto. La sección 3 introduce los fundamentos de las técnicas de virtualización, con especial énfasis en la herramienta VNUML, para posteriormente plantear en la sección 4 posibles escenarios de aplicación en el marco de los laboratorios de redes de la ETSIT-UPM. La sección 5 proporciona algunos detalles sobre la implementación de dichos laboratorios. Finalmente, se presentan las conclusiones en la sección 6.

2 Laboratorio de Redes de Ordenadores de la ETSIT-UPM

2.1 Objetivos y organización

Los objetivos docentes del laboratorio de redes de comunicaciones objeto de este artículo [1] se centran en las siguientes áreas:

- Estudio de protocolos de comunicación, su comportamiento y prestaciones con herramientas de simulación (NS –*Network Simulator*– [2]) y con analizadores de protocolos (*Ethereal* [3]).
- Diseño y planificación de redes. Esta área se basa en el uso de herramientas de simulación.
- Configuración de equipos de comunicaciones. Se dispone de una serie de entornos de red con infraestructura física instalada en el laboratorio,

que abarcan los siguientes ámbitos: LAN Ethernet, Frame Relay, ATM, RDSI, WLAN (WiFi). A estas subredes se conectan varios routers IP, de manera que con esta infraestructura se pueden instalar y configurar redes y servicios IP. Asimismo, los routers disponen de conexiones LAN a la que se conectan los puestos de laboratorio (PCs).
- Gestión y monitorización de red. Se dispone de una herramienta de gestión SNMP (*Simple Network Management Protocol*) comercial (HP OpenView), así como de diversas herramientas de monitorización y analizadores de protocolo.

Sobre estas áreas se organizan dos tipos de prácticas, según su organización:

- Prácticas de asistencia abierta al laboratorio, en la que los alumnos disponen de un plazo amplio para la realización de las prácticas, pudiendo acudir al laboratorio durante ese plazo, y disponiendo de la facilidad de reservas de puesto para la utilización de equipos concretos.
- Prácticas de asistencia controlada al laboratorio en horario concreto y limitado, en las que hay presencia del profesorado para la realización de prácticas monitorizadas o guiadas con asistencia de todo el grupo de alumnos.

Las primeras se adaptan mejor a las asignaturas de grado, permitiendo a los alumnos planificar la asistencia al laboratorio de acuerdo con sus necesidades. El segundo tipo se adapta más a los cursos de postgrado, que tienen un horario específico.

Con esta capacidad se da soporte a la docencia de varias materias, tanto de grado, como el Laboratorio de Ingeniería de Redes y Servicios Telemáticos (LRST) e Ingeniería de Redes y Servicios Telemáticos (IRST), como asignaturas de postgrado de los distintos cursos de maestría organizados por la ETSI Telecomunicación.

2.2 Escenarios de prácticas

Como ejemplo ilustrativo se describen las prácticas realizadas en la asignatura LRST. En cursos anteriores al 2004/2005 las prácticas se realizaban sobre bancos relativamente independientes. A partir de dicho curso, sin embargo, y con el objetivo de presentar a los alumnos un escenario único, se trabaja con el esquema mostrado en la Fig. 1, que pretende mostrar la posible organización de una red corporativa dotada de una sede central y diversas sedes regionales y sucursales. Las sucursales se conectan a sus sedes regionales mediante enlaces punto a punto, y las regionales a la sede central a través de una red troncal IP, con un enlace de respaldo basado en Frame Relay. Además, existen dos sedes adicionales conectadas mediante ATM.

La infraestructura está basada en equipamiento de diversos fabricantes: Cisco y Teldat para los routers, Fore y Virata para los conmutadores ATM, RAD para los conmutadores y multiplexores Frame Relay, y las distintas redes locales implantadas como VLANs sobre conmutadores Ethernet de HP y Cisco.

Este escenario general, complejo, evidentemente ofrece una gran versatilidad. Para cumplir con los objetivos de la asignatura LRST se realizan cinco prácticas, centrándose cada una de ellas en un subconjunto del escenario, o bien, si se considera globalmente, en el estudio de un protocolo o protocolos concretos. Así, se organizan en:

- Práctica de simulación. En ella se estudia el comportamiento del escenario, representado mediante un modelo de simulación, cuando se introducen protocolos no usados en la configuración habitual del laboratorio, concretamente soporte de calidad de servicio y protocolos multicast. Se dispone de la herramienta de simulación Network Simulator (NS), que dispone de bibliotecas con funciones de simulación de protocolos y redes de comunicaciones.
- Prácticas con equipos reales. Se organizan en:
  - Práctica de interconexión básica: configuración de la conectividad IP en el subconjunto formado por una sede regional y las sucursales dependientes.
  - Prácticas (2) sobre las subredes ATM y Frame Relay, basadas en los subconjuntos del escenario general donde aparecen estas subredes.
  - Práctica de encaminamiento dinámico con OSPF (*Open Shortest Path First*), sobre el escenario general.

Desde un punto de vista genérico, la práctica de simulación proporciona al alumno la oportunidad de experimentar tanto con escenarios de red como con tecnologías que o bien son difíciles de implantar en un laboratorio o bien suponen un gasto económico difícil de abordar. Por otro lado, las prácticas con equipos reales, consideradas globalmente, ofrecen al alumno la oportunidad de configurar escenarios de comunicaciones en los que entran en juego varias tecnologías de subred y diversos tipos de equipos (PCs, routers, conmutadores, servidores, etc.). Esto permite obtener una visión general de los principales protocolos y mecanismos involucrados en un escenario TCP/IP, asentando los conocimientos que sobre los mismos ya se han obtenido en otras asignaturas teóricas. Asimismo, el alumno se familiariza con diversos métodos de análisis y diagnóstico: utilidades específicas (ping, traceroute), mecanismos de trazado y depuración en los equipos de red, y captura y análisis de tráfico con analizadores de protocolos.

2.3 Evaluación del Laboratorio

Las evaluaciones realizadas entre el alumnado de cursos anteriores coinciden en valorar positivamente las actividades realizadas, destacando la alta puntuación que dan los alumnos a la posibilidad de

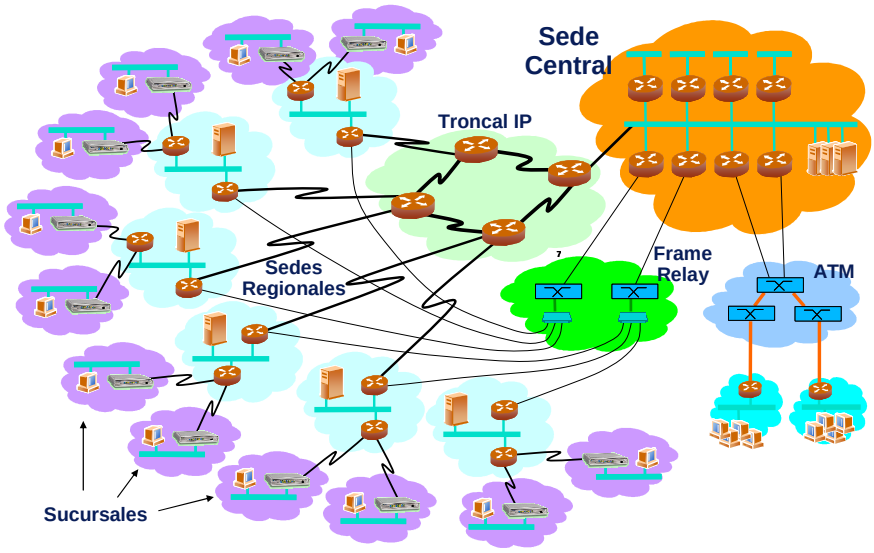


Figura 1: Escenario general del laboratorio

realizar prácticas sobre equipos reales, enfrentándose a problemas que no se aprecian desde la perspectiva mayormente teórica de otras asignaturas.

Así, la valoración desde el punto de vista docente es muy positiva. Sin embargo, desde el punto de vista operacional, la estructura del laboratorio presenta, junto a aspectos positivos como la gran variedad de posibilidades que ofrece la infraestructura, la dificultad que supone la operación y mantenimiento de un escenario general de estas características, así como la dificultad organizativa que presentan determinadas prácticas. Estas dificultades se pueden resumir en:

- Elevado coste del equipamiento involucrado, tanto para considerar posibles ampliaciones, como por mantenimiento y sustitución del equipamiento ya existente.
- La gestión del escenario consume muchos recursos. Si bien el uso de equipos como servidores de consolas y el uso de *scripts* de configuración simplifica la gestión, el coste global de la gestión sigue siendo importante.
- Determinados tipos de prácticas pueden exigir que todo o casi todo el escenario tenga que estar operativo (ej., la práctica de OSPF) para poder ofrecer resultados útiles desde el punto de vista docente. Esto implicaría que debe existir un número importante de alumnos realizando la práctica de manera coordinada, o bien que el esfuerzo exigido a los alumnos realizando la práctica en una sesión específica es muy grande.
- Planificación cuidadosa de los diferentes recursos que forman parte del escenario general, puesto que la infraestructura es compartida por diversas materias, y dentro de la misma materia distintas prácticas pueden requerir elementos comunes del escenario.
- Si bien la conectividad en el escenario general tiene la flexibilidad de que en gran parte es configurable (pertenencia a distintas VLANs – *Virtual Local Area Network*– [4]), algunas prácticas pueden requerir cambios físicos en el conexionado.

La realización de prácticas de simulación, si bien podría paliar algunas de estas dificultades, no puede plantearse como solución general, puesto que se pierde el objetivo de trabajo con equipos y configuraciones reales.

En este contexto surge una alternativa muy interesante en la posibilidad de *virtualización* de parte de la infraestructura del escenario. La virtualización supone la sustitución de infraestructura física por una infraestructura "virtual" que emula el comportamiento de la infraestructura física a la que sustituye, y que puede interaccionar perfectamente con la infraestructura real existente en el laboratorio. De esta manera se mantienen las ventajas de trabajar con equipamiento real, pues tanto los equipos reales como los equipos "virtualizados" ofrecen una interfaz real de configuración a los alumnos, así como la ejecución real, no simulada, de los protocolos de comunicación que se manejan en el escenario. Además, por tratarse de infraestructura virtualizada, se facilita enormemente la gestión y la configuración centralizada de un escenario complicado, consiguiendo paliar buena parte de las dificultades operacionales señaladas antes para un escenario de laboratorio basado exclusivamente en infraestructura física. En las siguientes secciones se detalla la técnica de virtualización (sección 3) que se propone para su uso en los laboratorios, y como puede ser utilizada (sección 4) e implementada (sección 5).

3 Técnicas de Virtualización

Desde el punto de vista genérico, podemos definir la *virtualización* como una técnica software que permite la ejecución de una *unidad de proceso* dentro de un entorno virtual que emula las condiciones de ejecución como si se tratara de un entorno real. La máquina física real que alberga el entorno de

ejecución -ejecutando el software de virtualización- se denomina *equipo anfitrión*. Actualmente, existen múltiples sistemas de virtualización, como por ejemplo el soporte de *jails* de FreeBSD [5] (virtualización de procesos), VMware [6] (virtualización de máquina hardware) y User Mode Linux [7] (virtualización de sistema operativo), sobre el que nos centraremos más adelante (sección 3.1).

La virtualización no se limita a la emulación de máquinas aisladas, sino que es posible (con la técnica adecuada) la creación de sistemas de "máquinas virtuales" que interaccionan entre ellas en un entorno que simula un escenario de red real, comportándose de forma equivalente al sistema real emulado (Fig. 2).

En el contexto de este artículo, el objetivo de la virtualización es facilitar la implementación de laboratorios docentes. Desde ese punto de vista, sus principales ventajas son el ahorro de costes de infraestructura y la simplificación de la gestión.

En primer lugar, el utilizar un único equipo (o unos pocos) para implementar toda una infraestructura (formada por múltiples equipos, incluyendo posiblemente la infraestructura de red para interconectarlos) supone un ahorro de costes, tanto mayor cuando más complejo (más equipos) contiene el sistema original. No solo hay que considerar el ahorro en hardware, sino también en espacio, un recurso a veces escaso en las universidades donde los laboratorios docentes son implementados.

En segundo lugar, la virtualización también aporta facilidades a la gestión. Un entorno formado por múltiples equipos interconectados es complejo de administrar, ya que las acciones de gestión (configuración, recuperación ante errores, actualización de software, etc.) involucran varios elementos. En el caso de utilizar virtualización, hay que actuar en un único punto (el equipo anfitrión) y los procedimientos de gestión se simplifican. Por ejemplo, en el caso de que la red de un host quede desconfigurada (una situación bastante habitual en el caso de laboratorios docentes donde los alumnos comenten errores durante el proceso natural de aprendizaje), un administrador podría restaurar el host desconfigurado simplemente recuperando un *backup* de su "disco duro virtual" (realizado en el momento de desplegar el laboratorio), ahorrándose el esfuerzo que supondría diagnosticar la causa del problema y arreglarlo sobre el host real en el caso convencional.

El principal problema al que se enfrentan las técnicas de virtualización es la transparencia (similitud con el sistema real emulado). El grado de similitud entre la implementación virtual del sistema y la realizada con equipos reales puede tomarse como una medida de la bondad de la técnica de virtualización empleada: lo deseable es que el sistema emulado se comporte lo más transparentemente posible, idealmente exactamente igual que el sistema real.

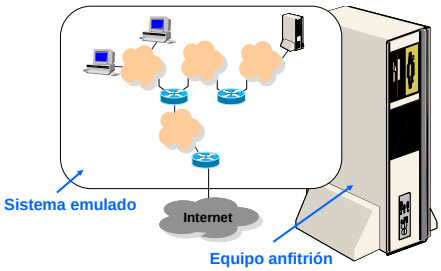


Figura 2: Ejemplo de sistema virtual

Si bien las técnicas de virtualización actuales permiten alcanzar gran transparencia en lo que se refiere a funcionalidad (es decir, los programas que se ejecutan en la sistema virtual se comportan igual que los que se ejecutan en el sistema real equivalente), es difícil alcanzar el mismo rendimiento (los programas en el sistema virtual se ejecutan "más lentos" que en el sistema real) por la sobrecarga de consumo de recursos debida a la virtualización. En el caso de laboratorios docentes orientados a funcionalidad, el menor rendimiento que supone la virtualización no suele suponer un problema.

3.1 User Mode Linux

UML (*User Mode Linux*) [7] es una modificación de las fuentes del núcleo de Linux que permiten su ejecución como proceso de usuario encima del núcleo convencional de Linux (el que ejecuta el equipo anfitrión). La funcionalidad del núcleo UML es exactamente la misma que la de un núcleo convencional de Linux, por lo que la transparencia en funcionalidad es completa. Cada proceso UML constituye una máquina virtual dentro de la cual otros procesos se ejecutan (Fig. 3).

Cada uno de los procesos UML tiene asociados sus propios recursos: espacio de memoria (multiplexando la memoria del equipo anfitrión), procesos (multiplexando la CPU del equipo anfitrión) y sistemas de ficheros (cada sistema de ficheros se implementa en un solo fichero en el equipo anfitrión). UML tiene optimizaciones (como el *copy-on-write*) que permiten ahorrar significativamente en el consumo de disco duro cuando existen múltiples máquinas virtuales.

UML no solo proporciona la manera de crear máquinas virtuales, sino los mecanismos para la interconexión entre ellas mediante redes virtuales (a nivel 2). Desde el punto de vista de la máquina virtual, se utiliza una interfaz de red de forma transparente, con la posibilidad de que la máquina anfitriona intervenga en la simulación o interconectar máquinas virtuales con equipos externos a través del interfaz físico, de forma completamente transparente o a través de VLANs. Una de las principales limitaciones de UML es que solo permite la creación de máquinas virtuales GNU/Linux. En todo caso, la emulación de plataformas de fabricantes o proveedores específicos no suele ser el objetivo en

laboratorios docentes, que normalmente persiguen la "neutralidad tecnológica" y la reducción de costes mediante el uso de sistemas libres.

En resumen, UML es una herramienta potente, sencilla de usar cuando se quiere utilizar una única máquina virtual, pero compleja si se pretende construir escenarios que incluyan muchas máquinas virtuales y topologías complicadas de red. Además, es necesario tener un buen conocimiento de ciertos detalles del sistema GNU/Linux (dispositivos *tap*, sockets UNIX, *bridges* virtuales, etc.) para construir escenarios "a mano".

3.2 La herramienta VNUML

La motivación con la que VNUML (*Virtual Network User Mode Linux*) [8] ha sido desarrollado es evitar la complejidad de uso de UML, de forma que el usuario se concentre en la definición del sistema emulado y no en los detalles propios de la virtualización. La herramienta ha demostrado su aplicabilidad a laboratorios docentes [9].

VNUML consta de dos componentes: lenguaje e intérprete. En primer lugar, el usuario describe el escenario a emular (máquinas virtuales, redes que las interconectan, configuración de los interfaces y rutas estáticas, etc.) en un lenguaje sencillo basado en XML (*eXtensible Markup Language*). A continuación, invoca un programa (*vnumlparser.pl*) que interpreta este lenguaje, creando el entorno virtual especificado. El intérprete interacciona con el sistema operativo del equipo anfitrión directamente, ejecutando las operaciones que el usuario tendría que hacer "a mano" en caso de no usar VNUML, ocultándole los detalles complejos.

El intérprete también permite automatizar la ejecución de grupos de comandos en las máquinas virtuales, lo cual es muy útil desde el punto de vista de la administración (por ejemplo, un grupo de comandos que permite devolver una práctica de laboratorio a su estado original). La Fig. 3 muestra un ejemplo simplificado del lenguaje de especificación de VNUML. Cada red virtual se especifica en una etiqueta *net* y cada máquina virtual en una etiqueta *vm* (dentro de la cual, cada interfaz se especifica con la etiqueta *if*). Como puede comprobarse, la sintaxis del lenguaje es tan sencilla que no requiere explicaciones adicionales.

4 Escenarios de utilización de VNUML

Tal como se ha mencionado en la sección anterior, las técnicas de virtualización y, en particular, la herramienta VNUML, tienen un campo de aplicación muy importante en la mejora de la gestión de un laboratorio de redes como el descrito en este artículo. Abren, además, la posibilidad de crear escenarios de red más complejos de lo que permitiría el equipamiento existente, permitiendo al alumno trabajar sobre escenarios de red más reales y mejorar con ello su aprendizaje.

Con el objeto de mostrar las distintas posibilidades que ofrece VNUML en el contexto general del laboratorio, se describen a continuación diferentes escenarios de utilización, en función de su grado de virtualización y de los sistemas a los que afecta. Aparte del nivel de virtualización 0, en el que todos los equipos son reales, podemos distinguir los siguientes escenarios.

4.1 Nivel 1: Virtualización de Elementos Auxiliares

En el escenario de la práctica, existen elementos auxiliares que facilitan la realización de la misma, aunque los alumnos no acceden de forma directa a ellos ni son objeto de configuración por su parte. Es el caso, por ejemplo, de los servidores de las sedes regionales y de la sede central (Fig. 1), que los alumnos utilizan simplemente para comprobar conectividad mediante el acceso a alguno de sus servicios (ping, web, ftp, etc). Es el caso también de los servidores de DNS, que facilitan que los alumnos puedan utilizar nombres IP en vez de direcciones a la hora de desarrollar las prácticas.

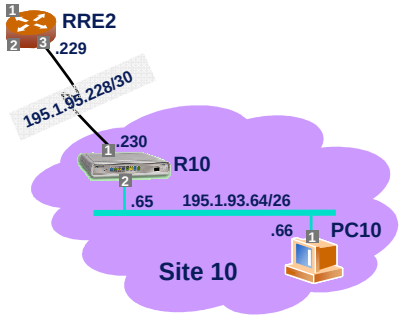
Estos elementos pueden ser implementados mediante VNUML como máquinas virtuales en un único equipo físico, reduciendo drásticamente el número de sistemas y el coste de administración de los mismos. Además, la utilización de servidores virtuales es transparente a los alumnos: por acceder a ellos de forma remota, no notarán diferencia alguna con respecto al mismo escenario basado en servidores reales.

Este enfoque, basado en la capacidad que proporciona VNUML para crear máquinas virtuales y conectarlas en cualquiera de las VLAN del laboratorio, proporciona un potencial enorme a la hora de completar escenarios de prácticas de laboratorio, permitiendo fácilmente añadirles todos aquellos servicios presentes en las redes de hoy en día. Además, permite particularizar la configuración de servicios para cada práctica o puesto de prácticas (cada uno puede tener su propio conjunto de servidores independiente), evitando interferencias.

Esta opción se utiliza regularmente en el laboratorio desde el curso 2003/2004, principalmente para crear los servidores de las sedes regionales y la sede central completa. Ello ha permitido eliminar los múltiples servidores físicos que se utilizaban anteriormente con este objetivo, reduciéndolos únicamente a uno.

4.2 Nivel 2: Virtualización de Elementos Principales

Algunas de las prácticas planteadas en el laboratorio necesitan para cumplir sus objetivos docentes que el




```
<vnuml>
...
<net name="link10" />
<net name="site10" />
...
<vm name="RRE2">
  <if id="1" ...>
  <if id="2" ...>
  <if id="3" net="link10">
    <ipv4
mask="255.255.255.252">195.1.95.229</ipv4>
  </if>
</vm>
```

```
<vm name="R10">
  <if id="1" net="link10">
    <ipv4 mask="255.255.255.252">195.1.95.230</ipv4>
  </if>
  <if id="2" net="site10">
    <ipv4 mask="255.255.255.192">195.1.93.65</ipv4>
  </if>
</vm>

<vm name="PC10">
  <if id="1" net="site10">
    <ipv4 mask="255.255.255.192">195.1.93.66</ipv4>
  </if>
</vm>
...
</vnuml>
```

Figura 3: Ejemplo de especificación en lenguaje VNUML

alumno las realice sobre el escenario completo de las mismas. Es el caso, por ejemplo, de la práctica sobre OSPF, que está pensada para ser realizada por 8/16 parejas de alumnos. Dicha práctica permite al alumno comprender el funcionamiento del protocolo de encaminamiento OSPF y analizar la composición de las tablas de encaminamiento de un router en función de los distintos parámetros de configuración (área única vs. múltiples áreas, inyección de rutas por defecto, agregación, etc). Es por ello que, para alcanzar los objetivos de la práctica, es necesario que estén funcionando un número mínimo de sucursales, de forma que las tablas de encaminamiento de los routers incluyan suficientes prefijos.

En un escenario sin virtualización, cada grupo de prácticas se ocupa de configurar una sucursal de la red y de configurar en colaboración con otro grupo su sede regional. Por tanto, es necesario un importante esfuerzo de organización previo a la realización de la práctica y de sincronización durante la misma para que todos los grupos vayan "al mismo ritmo". Por ejemplo, los router de las sedes regionales han de ser configurados coordinadamente entre dos grupos; si uno de ellos se retrasa en la realización de los ejercicios previos, provoca un retraso innecesario en sus compañeros.

La herramienta VNUML puede solventar en gran medida estos problemas, mediante la sustitución parcial o completa de los puestos de prácticas por escenarios virtuales. Es importante resaltar que los equipos a los que acceden directamente los alumnos siguen siendo reales. La virtualización en este caso se utiliza para implementar sistemas que no son objeto de configuración directa por parte de los alumnos.

4.2.1 Virtualización de puestos de prácticas no ocupados

Esta opción consiste en virtualizar los puestos de prácticas no ocupados, lo que permite que todos los alumnos realicen las prácticas en las mismas condiciones, independientemente del número de puestos ocupados en cada sesión. En el escenario del laboratorio descrito en este artículo, VNUML se utiliza para implementar sedes y sucursales completas en la práctica de OSPF (Fig. 1). De esta forma se elimina la necesidad de organizar las 8/16 parejas y la práctica puede realizarse con menos alumnos. Esta opción se está utilizando en la actualidad durante el curso 2004/2005.

En el caso extremo, puede virtualizarse todo el escenario salvo el puesto de prácticas de cada alumno, permitiendo así la realización de la práctica completa por un solo grupo y eliminar con ello la necesidad de coordinación entre grupos. Esta opción tiene la ventaja añadida de que permite realizar prácticas de gran complejidad a un coste comparativamente bajo, ya que el número de equipos reales requeridos es bajo.

4.2.2 Virtualización parcial del puesto de prácticas

En algunos casos es necesario reducir la complejidad o duración de una práctica, con el objeto de adaptarse a las condiciones particulares de cada curso y/o nivel del alumnado. Sin embargo, es importante que esta reducción no afecte sensiblemente a los objetivos docentes de la práctica y esta pueda ser realizada sobre el mismo escenario que la práctica completa.

Así, el trabajo del alumno puede concentrarse sobre una parte de los equipos del puesto de prácticas y utilizarse VNUML para emular el resto de los mismos. De esta forma, aunque el alumno trabaja solamente sobre una parte, los efectos observados son los mismos que en la práctica completa.

Por ejemplo, en el caso de la práctica de OSPF, si hay limitaciones de tiempo pero existe un grupo por cada sede regional, puede utilizarse esta opción para que los alumnos realicen una versión reducida de la misma. El trabajo del alumno puede concentrarse en la configuración de un único router (el router principal de la sede regional, que es el más interesante desde el punto de vista del encaminamiento) y proporcionarle como elementos virtuales el resto de equipos de las sucursales y sede regional. Esta opción ya ha sido utilizada con éxito durante el curso 2004/2005 en algunos laboratorios de programas de postgrado, en los que el tiempo es una limitación importante.

Por supuesto, si se dispone de infraestructura real para el escenario completo de la práctica, otra opción

consiste en la configuración por parte de los administradores/profesores de los equipos no configurados directamente por los alumnos. Sin embargo, el esfuerzo de despliegue y mantenimiento es mucho mayor, aun cuando la configuración de los equipos esté automatizada y no aporta ninguna ventaja funcional ni docente. Además, requiere la reserva y utilización exclusiva de dichos equipos, que en el caso de utilizar VNUML para emularlos quedan libres para otros usos, aumentando con ello la flexibilidad de organización del laboratorio.

Asimismo, VNUML facilita enormemente el cambio rápido de configuraciones, ya que permite definir de forma sencilla conjuntos de comandos a ejecutar sobre los sistemas virtuales, que pueden ser invocados en los momentos adecuados del guión de prácticas, según los alumnos avanzan en su realización (por ejemplo, para pasar de un escenario de área única a uno de áreas múltiples en OSPF).

4.2.3 Nivel 3: Virtualización Completa

La última de las opciones consiste en la creación de escenarios de prácticas completamente virtuales. Su principal ventaja es la gran reducción del coste que significa, ya que elimina la necesidad de adquirir equipos físicos reales (principalmente routers). Sin embargo, tiene la clara desventaja de que el alumno no utiliza equipos reales, lo que constituye un objetivo muy importante en laboratorios orientados a alumnos de últimos cursos o de postgrado.

Esta opción se puede aplicar a las siguientes situaciones:

- Laboratorios con escasa dotación económica, que no alcanza para construir escenarios reales, pero que sí disponen de recursos suficientes para construir un escenario virtualizado (básicamente ordenadores personales modernos con memoria por encima de los 512 Mb). De esta forma es posible, ofrecer prácticas de redes sobre escenarios de cierta entidad, con el mismo valor funcional que un escenario real.
- Laboratorios que disponiendo de equipamiento real, no les es posible dedicarlo en exclusiva a una sola práctica o a un grupo de alumnos. Sin embargo, en este caso las opciones más adecuadas serían las planteadas en 4.2.1 y 4.2.2.
- Realización del trabajo previo de una práctica. Muchas prácticas requieren que el alumno realice un estudio previo sobre la misma, de forma que cuando acceda al laboratorio tenga los conocimientos y habilidades necesarias para aprovechar eficientemente el tiempo asignado. En este sentido, pueden plantearse escenarios virtuales complementarios que le permitan experimentar con antelación aspectos importantes de la práctica que después verá en el laboratorio. Esta alternativa no se plantea como un sustituto de la práctica convencional, sino como un complemento, ya que la experiencia de configuración de equipos reales es importante.
- Realización de prácticas fuera del laboratorio (Auto-prácticas). El entorno de virtualización puede ofrecerse a los alumnos, junto con las configuraciones que permiten construir los escenarios experimentales, para que los alumnos puedan realizar las prácticas en sus propios equipos, sin las restricciones que supone la utilización de los laboratorios docentes (fundamentalmente relacionadas con horarios y el número de turnos de prácticas).

Dado que la herramienta VNUML puede ejecutarse casi en cualquier ordenador personal relativamente moderno, la última opción mencionada se plantea especialmente interesante para su utilización en cursos a distancia. Sería posible proporcionar a los alumnos un CD-ROM que contenga todos los elementos necesarios para que este realice la práctica fuera de horarios en un PC (por ejemplo, el PC doméstico del alumno) y de esa forma maximice su experiencia.

Además, para eliminar las dificultades de instalación, tanto del sistema operativo GNU/Linux como de VNUML, se plantea utilizar CDs autoarrancables, que permitan la ejecución inmediata de las prácticas. Esta opción está actualmente en desarrollo.

5 Realización del Laboratorio con soporte a la Virtualización

La Fig. 4 muestra el esquema general del laboratorio descrito en este artículo, incluyendo los servidores sobre los que se ejecutan los escenarios virtuales creados mediante la herramienta VNUML. La base del laboratorio está formada por un conjunto de conmutadores Ethernet interconectados entre ellos y a los que se conectan todos los interfaces Ethernet de los equipos de prácticas, principalmente los routers y los PCs. La configuración inteligente de las VLANs en los conmutadores permite crear fácilmente todas las subredes Ethernet que se necesitan en las prácticas y definir qué interfaces de los distintos equipos están conectados a ellas.

Se describen a continuación los componentes principales que forman el laboratorio:

- **Nodos de comunicaciones** (routers y conmutadores ATM y FR). Son los equipos principales de cada práctica y se localizan en los armarios del laboratorio. Los alumnos acceden a ellos mediante los servidores de consolas.
- **Ordenadores personales**. Son los equipos que utilizan los alumnos directamente y están dotados de dos placas Ethernet: la de producción, que permite el acceso a los servidores del laboratorio e Internet; y la experimental, que permite el acceso directo a los escenarios de cada una de las prácticas. Adicionalmente, algunos PCs llevan una tercera placa ATM o FR.

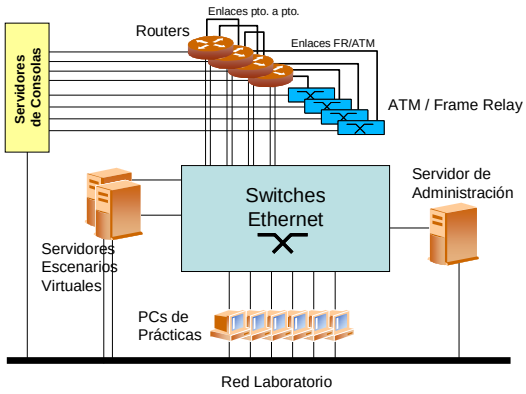


Figura 4: Esquema general de implementación

- **Servidores de Consolas**. Son equipos dotados de una conexión Ethernet y múltiples líneas serie. Permiten el acceso remoto a las consolas de los equipos de prácticas. De esta forma se evita la manipulación directa de los equipos y el consiguiente deterioro.
- **Servidor de Administración**. Este equipo es la pieza fundamental del laboratorio, ya que en el residen todos los *scripts* que permiten cargar las configuraciones iniciales necesarias para desarrollar las prácticas. Por ejemplo, los *scripts* que permiten reconfigurar mediante SNMP la VLAN a la que pertenece cada puerto de los conmutadores [11]; o los que cargan en los routers las configuraciones iniciales de cada práctica, basados en el lenguaje expect [10] y la utilización de un servidor de TFTP.
- **Servidores Escenarios Virtuales**. Son PCs estándar con 1 Gb de memoria. Se utilizan para ejecutar los escenarios virtuales necesarios en cada práctica. Por ahora, estos escenarios son controlados por los profesores, aunque está previsto que en el futuro puedan ser los alumnos quienes los arranquen directa o indirectamente a través, por ejemplo, de *scripts* desde páginas web. Las conexiones de estos servidores a los conmutadores Ethernet son del tipo "etiquetado", lo que permite que los equipos virtuales arrancados en un escenario virtual puedan ser localizados en cualquier VLAN de los escenarios de prácticas. Actualmente se usan dos servidores de este tipo, aunque no existe limitación para aumentar su número.

El desarrollo de *scripts* de administración que controlan los distintos equipos del laboratorio ha mejorado sensiblemente la gestión del mismo, permitiendo la reconfiguración completa del laboratorio el pocos minutos, sin necesidad de realizar las costosas configuraciones manuales, tan dadas a errores. Esto redunda en una utilización más eficiente de los recursos del laboratorio.

6 Conclusiones

El artículo propone la utilización de técnicas de virtualización en la implementación de laboratorios docentes. El uso de este tipo de técnicas, a distintos niveles, supone un importante ahorro en costes de espacio e infraestructura y, más importante, una simplificación en los procedimientos de organización y administración que, al final, repercute en el propio laboratorio y en la experiencia académica de los alumnos que lo utilizan. Es de destacar el uso de VNUML en los escenarios propuestos y la gran flexibilidad que la herramienta presenta. De hecho, sus casos de aplicación van más allá de los laboratorios docentes y es utilizada en otros contextos, como el diseño de prototipos para pruebas de aplicaciones de red [12]. Actualmente, VNUML sigue perfeccionándose e incorporando nuevas funciones, siendo los laboratorios docentes uno de los ejes impulsores de este desarrollo.

Referencias

[1] F.J. Ruiz, D. Fernández, et al, "Implantación de un laboratorio docente para redes de comunicaciones", JITEL 2001, Barcelona, Septiembre 2001.

[2] Página web de NS. Disponible en: http://www.isi.edu/nsnam/ns/

[3] Página web de Ethereal. Disponible en: http://www.ethereal.com

[4] "802.1q: Virtual LANs", IEEE 802.1Q Working Group, IEEE, 2001.

[5] P. Hope, "Using Jails in FreeBSD for Fun and Profit", Login: The Magazine of Usenix & Sage, n. 3, vol. 27, Junio 2002.

[6] J. Nieh and O. C. Leonard, "Examining VMware", Dr. Dobb's Journal, Agosto 2002.

[7] J. Dike, "User Mode Linux", Proc. 5th Annual Linux Showcase & Conf., Oakland CA, 2001.

[8] Página web de VNUML. Disponible en: http://www.dit.upm.es/vnuml

[9] F. Galán, D. Fernández, F. J. Ruiz, O. Walid, T. de Miguel. "Use of Virtualization Tools in Computer Network Laboratories", ITHET'04, Estambul, Mayo 2004. ISBN: 0-7803-8596-9. IEEE Catalog Number: 04EX898.

[10] Don Libes, "Exploring Expect: A Tcl-Based Toolkit for Automating Interactive Programs", O'Reilly and Associates, 1995.

[11] S. Fernández, "Implementación de un Sistema de Gestión de Red para la Red Gigabit de la ETSIT", PFC, Julio 2004.

[12] D. Fernández, F. Galán, T. de Miguel, "Study and Emulation of IPv6 Internet Exchange (IX) based Addressing Models", IEEE Communications Magazine, vol. 42(1), pp. 105-112, Enero 2004. ISSN: 0163-6804.

Edukalibre: Colaboración para la elaboración de material didáctico

Diego Chaparro, Luis López, Jesús M. González-Barahona,
{dchaparro,llopez,jgb}@gsyc.escet.urjc.es
Grupo de Sistemas y Comunicaciones (GSyC). Universidad Rey Juan Carlos
Escuela Superior de Ciencias Experimentales y Tecnología.
C/ Tulipán s/n. 28933 - Móstoles (Madrid)

Abstract *The Libre Software collaborative development model has proved to be a powerful tool for the creation of software at many different scales and on multiple application domains. In this paper, we propose a system developed in the framework of the Edukalibre project, which intends to import and adapt this model for the creation and management of educational contents. This system supplies a set of tools for the creation of collaborative contents allowing version controlling, automatic conversions between multiple formats, diverse accessing methods and tools, etc. A fully functional version of the system is already available to download under a Libre Software license.*

1. Introducción

Las tecnologías web se han introducido en nuestro mundo de manera rápida invadiendo múltiples aspectos de nuestras vidas. Esta revolución está teniendo un impacto notable sobre el modo en que trabajamos, nos relacionamos y aprendemos. En la actualidad, es posible acceder a multitud de diferentes tipos de contenidos en cualquier instante y desde cualquier lugar. Por este motivo, numerosas instituciones, organizaciones y empresas del mundo de la educación están explorando las posibilidades de aplicar estas tecnologías de la información en sus procesos educativos [1, 16, 7].

En la actualidad, la mayoría de estos sistemas se utilizan como complemento a las técnicas tradicionales de enseñanza, por lo que se basan en el modelo tradicional profesor/instructor frente al de alumno/estudiante. Por este motivo, la mayor parte de los contenidos educacionales que se pueden encontrar en la web han sido diseñados para un modelo de tele-enseñanza a través de tele-lectula y consisten en estructuras monolíticas y fijas, secuenciadas y limitadas que no están dotadas de flexibiliad para la realización de actualizaciones y que son difíciles de organizar y reutilizar [18].

Debido a esto, existe un gran interés en la comunidades científicas y académicas por concebir nuevos modelos y metodologías que permitan disponer de mecanismos telemáticos eficientes para la creación, compartición y reutilización de recursos educacionales en la web. En particular, se ha detectado la necesidad de incidir en nuevos paradigmas que posibiliten el e-learning basándose en modelos más colaborativos y proactivos.

En este artículo proponemos un sistema de tele enseñanza que desarrolla estas ideas utilizando para ello un conjunto de metodologías y técnicas derivadas de las que se usan en el desarrollo de Software Libre. Estos métodos han modificado enormemente el modo en que el software se produce y se distribuye [2, 6]. La idea fundamental se basa en la presencia de una comunidad de participantes que comparten experiencias, conocimiento y código fuente en un proceso colaborativo en el que cada individuo ayuda con lo que puede. Ha sido aplicada con éxito en múltiples dominios incluyendo el desarrollo de sistemas operativos (Debian, FreeBSD, Fedora), entornos de escritorio (GNOME, KDE), navegadores y servidores web (Mozilla, Firefox, Apache) o aplicaciones ofimáticas (OpenOffice.org) [4, 5, 15, 11]. Por este motivo, hoy en día está aceptado que las comunidades del software libre han producido métodos revolucionarios en el ámbito de la ingeniería del software [14, 10].

Es importante destacar que existen multitud de similitudes entre el modelo de desarrollo del software libre y algunas metodologías modernas de aprendizaje basadas en la cooperación. De hecho, algunos autores proponen explícitamente que las actividades que tienen lugar en las comunidades del software libre pueden ser vistas como un proceso de aprendizaje en sí mismo, en el que las partes involucradas contribuyen y aprenden del resto [3]. Consiguientemente, es de esperar que la aplicación de dicho modelo puede tener un importante impacto sobre el modo en que las tecnologías web se utilizan para enseñar y aprender. Por este motivo, creemos que es extremadamente importante que las comunidades académicas y científicas que desarrollan o utilizan sistemas telemáticos para la enseñanza conozcan las potencialidades de este modelo.

En la actualidad, existe un creciente interés por importar algunos aspectos de los modelos de desa-

rrollo del software libre en los procesos de aprendizaje. Los primeros pasos los han dado algunas instituciones educativas de gran prestigio tales como el MIT [12], la Carnegie Melon University [8] o Harvard [13]. Estas universidades han sobrepasado el esquema tradicional de pensamiento (de acuerdo con el que los materiales sólo están disponibles para los alumnos matriculados en los cursos correspondientes) y proporcionan un acceso completamente libre a su material educativo de alta calidad a través de la web. Algunos aspectos económicos o tecnológicos pueden haber influido en esta decisión, incluyendo beneficios de márketing, incremento de la reputación, mayor difusión de la innovación en la sociedad, posibilidad de obtener realimentación de profesionales y estudiantes de todo el mundo, etc. Esta idea de abrir los contenidos educativos se está desarrollando con gran entusiasmo en un importante número de proyectos entre los que se encuentran MIT OpenCourseware [12], Open Learning Initiative [8], etc.

La idea de hacer los contenidos universalmente accesible podría producir una revolución en la educación similar a la que se produjo en el mundo del software cuando se introdujo el software libre. Sin embargo, para que esta revolución realmente pueda tener lugar, simplemente hacer que los contenidos educativos sean libremente accesibles no es suficiente. El software libre no puede ser comprendido exclusivamente en función de la accesibilidad del código fuente. Su éxito se basa también en la existencia de mecanismos por los que los diferentes agentes se coordinan y obtienen un beneficio a través de la cooperación. El proceso por el que esta comunidad surge y se forman sinergias es complejo y no se comprende plenamente. Sin embargo, nuestra hipótesis es que suministrando facilidades similares a las que están disponibles para el desarrollo de código, se podría producir un fenómeno análogo en el ámbito de la generación de contenidos educativos. Es de destacar que esta hipótesis es razonable puesto que algunas iniciativas exitosas tales como Wikipedia [19] están basadas en la misma filosofía.

La idea de producir aprendizaje colaborativo en una comunidad no es nueva. Algunos autores [17, 9] describen una comunidad dinámica de aprendizaje como aquella en la que el control se distribuye entre miembros autónomos que pueden involucrarse de manera flexible en actividades de aprendizaje de modo negociado y dialogado, interactuando y colaborando con un claro compromiso hacia la generación y compartición de nuevos conocimientos. En estos trabajos, se reconoce que, a través de la promoción de la creatividad, la innovación, la colaboración y proporcionando mecanismos que permitan la modificación de contenidos, la comunidad es capaz de diagnosticar y solucionar la mayoría de las necesidades educativas que puedan surgir. Sin embargo, no se establece de manera clara cómo es posible alcanzar estos objetivos

Nuestra propuesta es la de integrar la filosofía del software libre en la educación para lograrlo. Hasta hace poco, el estado de la tecnología hacía muy difícil soportar el desarrollo de contenidos educativos construidos de manera colaborativa entre grandes grupos de profesores y alumnos. La comunidad del software libre ha creado multitud de herramientas que permiten que la colaboración pueda llevarse a cabo entre personas en diferentes lugares e instantes. Sin embargo, estas tecnologías no son fácilmente aplicables en el entorno educativo debido a diversos motivos. En primer lugar, no son lo suficientemente intuitivas para ser utilizadas por profesores y estudiantes medios. En segundo, están diseñadas para realizar tareas parciales, lo que se adapta al perfil de los desarrolladores que conocen y pueden utilizar múltiples herramientas complementarias, pero que no es satisfactorio en comunidades educativas.

En este contexto, está claro que son necesarias nuevas herramientas que permitan realizar de manera efectiva el desarrollo colaborativo de ecursos educacionales abiertos. Una solución basada en la web parece ser la elección más razonable para implementarlas por dos motivos. En primer lugar, esta tecnología ofrece la posibilidad de integrar todo tipo de contenidos y formatos. En segundo, está muy extendida y puede ser utilizada por un usuario medio sin ningún tipo de formación especializada. Más aun, estas herramientas deberían ser también softwar libre para garantizar que puedan ser adaptadas y desplegadas utilizando el mismo concepto abierto de colaboración. En este artículo, proponemos una aplicación que cumple todos esos requisitos y que permite la concepción de contenidos educacionales colaborativos en la web. Esta aplicación ha sido desarrollada dentro del marco del proyecto Edukalibre, financiado por la Comisión Europea bajo el programa Socrates/Minerva [1]. El proyecto comenzó en Octubre de 2003 y se espera que termine en Diciembre de 2005. Es coordinado por la Universidad Rey Juan Carlos e incluye como partners a la Universidad de Leeds, la Universidad de Porto, y la Universidad de Karlsruhe. La página web de proyecto se encuentra en http://www.edukalibre.org.

El Sistema Edukalibre se describe en el resto de este artículo. En primer lugar describiremos la arquitectura básica del mismo. Acto seguido se presentarán los diferentes módulos que implementan cada una de las funcionalidades, así como las interfaces a través de las que interactúan. Finalmente se expondrán unas breves conclusiones.

[1] http://europa.eu.int/comm/education/programmes/socrates/minerva/ind1a_en.html

2. Arquitectura modular del sistema

La arquitectura de la plataforma edukalibre se basa en un sistema de módulos intercambiables, de modo que cada uno de ellos puede ser substituido por otro de características similares sin necesidad de modificar el resto. También es posible añadir nuevos módulos de forma sencilla, permitiendo que estos se integren con el resto a través de una interfaz predefinida.

Los módulos se agrupan en varios niveles, cada uno de los cuales debe proporcionar una funcionalidad preestablecida. En la actualidad, el sistema está constuído en base a dos niveles. El nivel inferior, sobre el que se sustenta el resto de la aplicación, se basa en un sistema de control de versiones de ficheros. Este proporciona las funcionalidades básicas de gestión de la historia de los documentos. El nivel superior está formado por varios interfaces de usuario que permiten realizar diferentes tareas, que van desde las más básicas y habituales hasta las más avanzadas restringidas a usuarios expertos. La figura 2 muestra un esquema de la arquitectura y la interrelación entre los módulos.

La materia prima de la que se nutre el sistema son los documentos. Por tanto, para comprender en detalle cómo interaccionan los diferentes componentes de la aplicación, hay que comprender con precisión qué entendemos por un documento y cómo estos se gestionan. La sección siguiente está dedicada a este objetivo.

2.1. La noción de documento

El documento es la unidad funcional en la que se basa el sistema. En ellos, los usuarios pueden introducir y recuperar información. Para esto, se ofrecen un conjunto de operaciones que permiten manipularlos. Los documentos se pueden clasificar en varios tipos dependiendo de su formato

- *Documentos base*: son documentos que pueden ser editados por los usuarios para crear un nuevo documento, o generar una nueva versión de un documento preexistente. Para que un documento base sea aceptado por el sistema, su formato debe pertenecer al conjunto de formatos base. Los elementos de dicho conjunto han sido seleccionados gracias a sus propiedades y ventajas. En la actualidad se soportan los siguientes: DocBook, LaTeX y OpenOffice. Los dos primeros se representan mediante texto plano (no son ficheros binarios) lo que proporciona ventajas a la hora de almacenarlos en un sistema de control de versiones y también para extraer información de los mismos. El tercero, aunque en principio no se representa como un fichero de texto plano, en realidad es un fichero comprimido que contiene, entre otras cosas, varios ficheros XML con toda la información del documento. Por tanto, también puede procesarse de manera similar a los dos anteriores.

- *Documentos finales*: Los documentos finales son aquellos cuyo formato se obtiene a partir de los documentos base. Una característica destacable del sistema es que se generan conversiones de los documentos base a los finales automáticamente cada vez que se introduce un nuevo documento o se genera una nueva versión de uno existente. Los formatos para los documentos finales se han elegido para facilitar la distribución y el uso de los mismos. Por ejemplo, algunos de los que se generan actualmente son:

  - Aptos para ser impresos:
    - Postcript
    - PDF
  - Aptos para ser publicados en la web:
    - HTML múltiple
    - HTML único
  - Totalmente portables:
    - Texto plano

Es importante destacar que el sistema Edukalibre proporciona un nivel de abstracción superior al de otros sistemas similares debido a que la unidad de trabajo es el documento y no el simple fichero. En este sentido, un documento puede constar de múltiples ficheros que pueden contener el texto del documento, imágenes, hojas de estilo para realizar las conversiones de formatos, etc.

2.2. Nivel 0: Núcleo del sistema

Este nivel está formado por un conjunto de módulos que proporcionan la funcionalidad principal al sistema. En él se almacenan todos los documentos, tanto los que están en formato base como los que se encuentran en formatos finales. También en este nivel se realizan las conversiones entre estos formatos y se extrae automáticamente un conjunto de informaciones que son relevantes para las capas superiores a la hora de interaccionar con los usuarios (p.e. título, autores, etc.) Cada una de estas funcionalidades es implementada por uno o varios módulos específicos. A continuación explicaremos en detalle en qué consiste cada uno de ellos.

2.2.1. Módulo: Repositorio de documentos base

Es el módulo principal de este nivel, y está formado básicamente por un sistema de control de

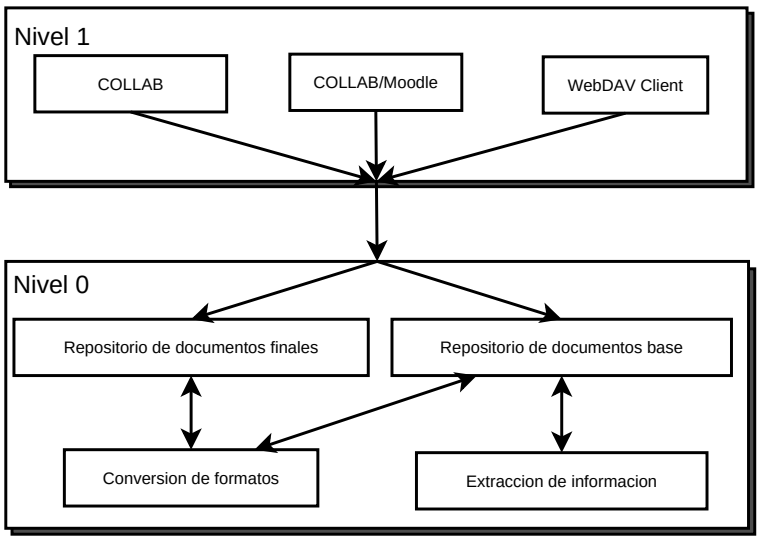


Figura 1: Arquitectura del sistema

versiones. En la versión actual se utiliza subversion[2] pero el sistema está preparado para adaptarse fácilmente a otro tipo de soluciones tales como CVS[3] o GNUArch[4]. El repositorio es capaz de almacenar un registro que contiene todas y cada una de las diferentes versiones de cada uno de los documentos que se hayan insertado en el sistema. A través de su interfaz, se pueden obtener los datos necesarios para poder llevar a cabo tareas tales como recuperar estados anteriores, hallar las diferencias entre dos versiones, analizar el grado de contribución de los diferentes autores y determinar en qué partes del documento se ha producido dicha contribución, etc.

Además, el sistema de control de versiones implementa un mecanismo que permite añadir pequeños programas que se ejecutan cuando se realiza una determinada tarea sobre el repositorio. En la actualidad, estos programas se dividen en dos fases coincidiendo con dos de las operaciones que realiza subversion: el "pre-commitz el "post-commit".

- El conjunto de tareas asociadas a la *operación "pre-commit"* se ejecutan cuando se trata de subir un documento, pero éste no ha sido comprometido aún en el repositorio. En esta fase, se efectúan tareas de validación de documentos, que consisten básicamente en la comprobación sintáctica de los documentos base que se intentan introducir en el sistema. Por ejemplo, en el caso de documentos DocBook, la validación se producirá contra la DTD correspondiente. Esta tarea es importante, porque si un documento no respeta la sintaxis correcta de su especificación, hay otras tareas posteriores que no funcionarán correctamente como la conversión automática de formatos. En esta fase, se puede actuar de dos formas cuando un documento no es válido, o se aborta la operación mostrando el mensaje de error al usuario, o se continúa la operación y al final se le muestra un mensaje de aviso al usuario indicándole los posibles problemas.

- La segunda fase está asociada a la *operación "post-commit"* del repositorio, la cual ejecuta el código asociado cuando el documento ya se ha subido al sistema y ya no se puede abortar esta operación. Esta segunda fase se encarga de controlar el flujo de control del resto de módulos de este nivel, gestionando la ejecución de los mismos. Primero se ocupa de extraer información sobre el documento que ya se ha comprometido en el repositorio, y después activa el módulo encargado de realizar las conversiones de formatos sobre ese documento.

2.2.2. Módulo: Repositorio de documentos finales

Además del repositorio base, existe otro repositorio destinado a almacenar los documentos en sus formatos finales. Éste último se encuentra integrado con un servidor HTTP, que en la versión actual es Apache 1.3. De este modo, es posible acceder mediante este protocolo a toda la historia de cada uno de los documentos en su formato final.

El motivo por el que se ha separado el almacenamiento de documentos base y finales en dos repositorios diferentes es claro. Los documentos base son los únicos susceptibles de ser editados por los

[2] http://subversion.tigris.org
[3] https://www.cvshome.org/
[4] http://www.gnuarch.org

usuarios. Ya hemos visto que los formatos base se basan todos en ficheros de texto, lo que permite utilizar plenamente todas las funcionalidades de los sistemas de control de versiones. Por el contrario, los documentos finales suelen ser ficheros binarios, que el usuario final no puede modificar y que cambian de manera automática cuando sus correspondientes bases son alteradas. En este caso, es suficiente utilizar un servidor HTTP que permita descargar directamente los correspondientes ficheros bajo demanda. Por estos motivos, las funcionalidades que se requieren en ambos casos por parte del repositorio son diferentes y parece lógico aislar ambas alternativas.

2.2.3. Módulo: Conversión de formatos

La conversión de formatos es un módulo clave de este nivel, dado que genera directamente el producto que esperan los usuarios a través de un mecanismo automático que se activa cada vez que un nuevo documento o una nueva versión se introducen en el sistema.

El diseño de este módulo ha sido cuidadosamente estudiado, debido a que los procesos que se llevan a cabo en el mismo pueden suponer una carga computacional elevada. Dependiendo de la complejidad de los documentos que se traten, es posible encontrar que ciertas conversiones de formatos pueden llegar a tomar un tiempo respetable para poder llevarse a cabo. Este factor ha de ser tenido en cuenta para que no interfiera en la interacción con el usuario.

Las conversiones desde el formato base DocBook se realizan utilizando un procesador XSLT, para lo que se ha diseñado una hoja de estilo para cada uno de los formatos que generamos. Los procesadores de hojas de estilo utilizados son *xsltproc* y *fop*. Usando este mecanismo, se generan los siguientes formatos finales: HTML único, HTML múltiple, texto plano, PDF, Postcript, LaTeX y OpenOffice. Algunas de las hojas de estilo XSL usadas están disponibles desde hace tiempo para los procesadores que hemos citado aunque, en algunos casos, estas se han modificado para adaptarlas a nuestras necesidades.

Las conversiones desde el formato base LaTeX se realizan utilizando las herramientas clásicas disponibles en el mundo del software libre para realizar conversiones desde LaTeX: *latex*, *dvipdf*, *dvips* y *hevea*. Con estas herramientas conseguimos los siguientes formatos finales: DVI, HTML único, HTML múltiple, PDF y Postcript.

Y las conversiones del formato base OpenOffice se realizan utilizando el propio editor OpenOffice.org en modo comando. Los formatos finales que es posible obtener en este caso son los siguientes: HTML único, PDF, Microsoft Word y texto plano.

Evidentemente, hemos descrito solamente una pequeña muestra de la gran colección de conversiones que se podrían realizar de forma sencilla utilizando herramientas ya disponibles desde cada uno de los formatos base. El sistema permite añadir fácilmente nuevos formatos finales que se generan desde los base, e incluso formatos finales que se generan desde otros formatos finales. En la actualidad seguimos trabajando en ampliar la lista de formatos finales para cada uno de los formatos base lo máximo posible, de modo que se puedan satisfacer las necesidades de todos los usuarios.

2.2.4. Módulo: Extracción de información de documentos

Otra característica interesante del sistema es que proporciona a los usuarios información acerca de cada documento sin necesidad de tener que acceder al propio documento. Para ello, se extraen ciertos campos de cada documento a partir de los formatos base que se introducen en el sistema. Actualmente este proceso sólo se realiza sobre los documentos base de tipo DocBook/XML, que se someten a un parseado en el que se localizan las etiquetas necesarias en el fichero. La implementación de esta funcionalidad en los otros formatos base está lista en breve.

La información que se extrae actualmente de los documentos es la siguiente:

- Título
- Abstract
- Autor

Esta información se almacena en el repositorio de documentos finales, junto con otros datos interesantes como la fecha y hora en la que se ha subido el documento al sistema, el usuario que lo ha subido, el nombre de la revisión del documento, etc.

2.3. Nivel 1: Interfaces del sistema

El sistema puede ser accedido por los usuarios desde diferentes interfaces y a través de diferentes protocolos de comunicaciones. Hemos desarrollado algunas interfaces de forma explícita para ser usadas sobre nuestro sistema; entre ellas COLLAB y COLLAB/Moodle que explicamos en detalle en esta sección. Sin embargo, también es posible utilizar interfaces genéricas y estandarizadas como el protocolo WebDAV, que permite acceder al repositorio de documentos base, o el HTTP que proporciona el repositorio de documentos finales.

De este modo, proporcionamos a los usuarios diversas alternativas a la hora de acceder al sistema. Por un lado, los usuarios medios pueden descargar los documentos finales utilizando navegadores web convencionales o pueden utilizar un cliente WebDAV para realizar tareas sencillas sobre los base. Por otro lado, existen mecanismos de acceso que requieren un nivel de conocimiento más técnico, como el interfaz que proporciona Subversion con sus herramientas de cliente tradicionales, que

pueden ser usadas como con un servidor de Subversion tradicional por usuarios más avanzados.

2.3.1. Módulo: COLLAB

Este es uno de los interfaces de usuario que hemos desarrollado para acceder al sistema y que permite a los usuarios realizar las tareas más comunes sobre el sistema Edukalibre de una forma sencilla.

Básicamente COLLAB es un interfaz web escrito en PHP que puede ser integrado de modo sencillo en cualquier portal escrito en este lenguaje. COLLAB proporciona la funcionalidad suficiente para que cualquier usuario pueda realizar las acciones más comunes sobre el sistema, tal y como se muestra a continuación. Los usuarios avanzados que requieran funcionalidad no contemplada en este interfaz, pueden usar otras herramientas más genéricas como se muestra en el apartado 2.3.3.

- Autenticación: COLLAB soporta un método de autenticación para acceder a ciertos aspectos de su funcionalidad restringida a usuarios registrados
- Acceso a documentos: mediante este interfaz podemos ver los documentos almacenados en el sistema e información sobre los mismos.
- Historia de documentos: se puede observar la historia de cada documento, mostrando cada una de las versiones, con su fecha, la persona que la subió al sistema, y enlaces para acceder a cada versión del documento en su formato base o en los finales.
- Introducir nuevos documentos o actualizar documentos existentes: mediante este interfaz se puede subir al sistema un nuevo documento o una nueva versión de un documento existente utilizando simplemente un formulario web.
- Edición on-line: Proporciona un sencillo interfaz para poder editar mediante el navegador los documentos (en formato base DocBook o LaTeX).

Puede verse la apariencia del interfaz COLLAB en la Figura 2 .

2.3.2. Módulo: COLLAB/Moodle

El uso de los Sistemas de Gestión de Contenidos (CMS) se está extendiendo por muchos centros educativos en la actualidad. Por un lado es utilizado para gestionar cursos de enseñanza a distancia, y por otro lado como apoyo a la gestión de los cursos presenciales tradicionales. Moodle es uno de los CMS que se está extendiendo en la comunidad educativa gracias a sus ventajas y funcionalidades y también gracias a la licencia GPL con la que se distribuye.

COLLAB/Moodle es una adaptación del interfaz COLLAB que puede ser usado como un recurso dentro de los cursos de Moodle. Se ha desarrollado de forma que pueda ser instalado de forma sencilla, ya que se empaqueta como un módulo Moodle y está totalmente integrado con el CMS tanto en funcionalidad y gestión de la base de datos, como en la apariencia, gracias a las hojas de estilo de Moodle.

En la Figura 2 puede verse la apariencia del interfaz COLLAB/Moodle.

2.3.3. Interfaces genéricas

Uno de los objetivos perseguidos en el diseño del sistema es que pueda ser accesible desde el mayor número de herramientas que sea posible. Para lograrlo, además de implementar varios interfaces de usuario como COLLAB y COLLAB/Moodle, el sistema soporla la utilización de protocolos de comunicaciones estándar.

De este modo, los repositorios pueden ser accedidos usando los protocolos HTTP (Hyper Text Transfer Protocol), y WebDAV (Web-based Distributed Authoring and Versioning) que es una extensión de HTTP que permite editar y gestionar ficheros en servidores remotos.

- *Accesos mediante HTTP*: Se puede acceder al sistema mediante el protocolo HTTP directamente sobre el repositorio de documentos base y sobre el repositorio de documentos finales. Mediante este acceso podemos realizar tareas de lectura sobre estos repositorios; es decir, podemos acceder a los documentos base y finales, pero solo en modo lectura. No podemos modificar los documentos ni subir otros nuevos.
- *Accesos mediante WebDAV*: Mediante el acceso WebDAV podemos realizar las mismas acciones que utilizando HTTP. Sin embargo, en estecaso, además es posible llevar a cabo otras operaciones utilizando las ampliaciones que proporciona WebDAV. Así, es posible acceder a propiedades de los ficheros (método PROPFIND), crear directorios (método MKCOL), copiar y mover ficheros (métodos COPY y MOVE), acceder a la historia de un documento o mezclar varias versiones de un documento.

Para acceder mediante el protocolo HTTP, podemos usar cualquier cliente que lo soporte, incluyendo navegadores web (Mozilla, Netscape, etc.) o editores que permitan leer los documentos desde un servidor HTTP (casi cualquier editor lo permite), pero recordando que sólo podemos acceder en modo lectura, es decir, no podremos guardar el documento en el repositorio usando este protocolo.

El acceso mediante WebDAV se puede realizar mediante un explorador de ficheros que lo soporte

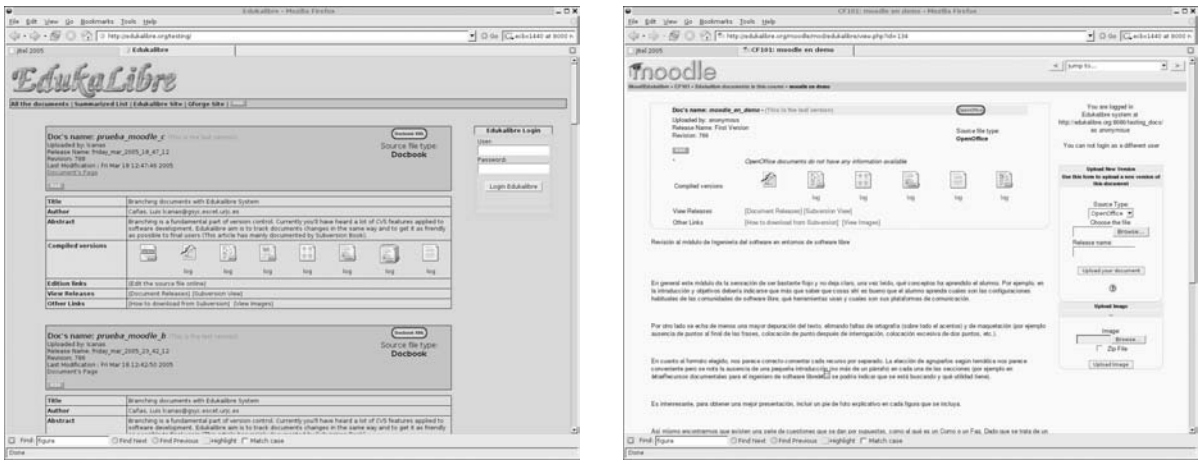


Figura 2: Página principal de COLLAB e interfaz de COLLAB/Moodle

(como Nautilus en GNOME o cadaver[5] que es un cliente WebDAV en modo texto). También puede realizarse utilizando directamente un editor compatible como OpenOffice.org[6] o emacs[7].

Además, el sistema tiene disponible un interfaz del tradicional ViewCVS, [8] que presenta mediante un interfaz web información sobre los documentos almacenados en el repositorio de documentos base, mediante el que se pueden mostrar diferencias entre versiones de un fichero, historia de versiones de un fichero, etc. Una de las diferencias entre este interfaz y el interfaz que proporciona COLLAB es que el primero está centrado en ficheros, mientras que el segundo trabaja sobre la abstracción de documentos, que a su vez pueden estar compuestos de uno o varios ficheros.

3. Conclusiones

In este artículo hemos descrito el Sistema Edukalibre, una nueva aplicación que permite el desarrollo colaborativo de contenidos educativos y que se basa en las metodologías de desarrollo de software libre. Hemos presentado las principales características del sistema así como su arquitectura básica. Hemos mostrado que este tipo de soluciones satisface plenamente las necesidades que pueden surgir a la hora de crear este tipo de contenidos. Por este motivo, proponemos que la herramienta puede tener aplicaciones interesantes dentro del ámbito de la tele-educación, por lo que creemos que la comunidad de la Ingeniería Telemática debe conocer su existencia. El sistema ha sido desarrollado utilizando herramientas libres y ha sido liberado bajo una licencia también libre. Por tanto, puede ser descargado, utilizado y distribuido con todas las ventajas asociadas a este tipo de modelo.

[5] http://webdav.org/cadaver/
[6] http://www.openoffice.org/
[7] http://www.gnu.org/software/emacs/
[8] http://viewcvs.sourceforge.net/

Agradecimientos

El trabajo descrito en este artículo ha sido financiado en parte por el programa Socrates/Minerva de la la Comisión Europea, bajo el número de ayuda 110330-CP-1-2003-1ES-MINERVA-M.

Referencias

[1] H.H. Adelsberger, B. Collis, and J.M. Pawlowski. *Handbook on Information Technologies for Education and Training.* Springer, 2002.

[2] Nikolai Bezroukov. Open source software development as a special type of academic research. *First Monday*, 4(10), October 1999.

[3] K. Edwards. Epistemic communities, situated learning and open source software development. http://opensource.mit.edu/papers/kasperedwards-ec.pdf.

[4] Daniel M. German. The GNOME project: a case study of open source, global software development. *Software Process Improvement and Practice*, pages 201–215, August 2003.

[5] Jesús M. González-Barahona, Luis López, and Gregorio Robles. Community structure of modules in the apache project. In *Proceedings of the 4th Workshop on Open Source Software Engineering. 26th International Conference on Software Engineering*, Edinburgh, Scotland, UK, May 2004.

[6] Ahmed E. Hassan, Michael W. Godfrey, and Richard C. Holt. Software engineering research in the bazaar. In *Proceedings of the 2nd Workshop on Open Source Software Engineering at the 24th International Conference on Software Engineering*, May 2001.

[7] W. Horton. *Designing Web-Based Training.* Wiley Computer Publishing, 2000.

[8] Carnegie Mellon Open Learning Initiative. http://www.cmu.edu/oli/.

[9] P. Irvine and P. Brna. Growing an internet-based community for liefelong self-learners: empowering the community. *International Journal of Continuing Engineering Education and Lifelong Learning*, 13(1/2):21–21, 2003.

[10] Stefan Koch, editor. *Free/Open Source Software Development.* Idea Group, Inc., 2004.

[11] Christoph Lameter. Debian gnu/linux: The past, the present and the future. In *Free Software Symposium 2002*, October 2002.

[12] MIT OpenCourseWare. http://ocw.mit.edu/index.html.

[13] Harvard University Library Open Collections Program. http://ocp.hul.harvard.edu.

[14] Eric S. Raymond. The cathedral and the bazar. *First Monday*, 1997. http://www.firstmonday.dk/issues/issue3_3/raymond/.

[15] Christian Robottom Reis and Renata Pontin de Mattos Fortes. An overview of the software engineering process and tools in the Mozilla project. In *Workshop on Open Source Software Development*, February 2002.

[16] J.M. Rosenberg. *E-Learning.* McGraw-Hill, 2001.

[17] M. Simonson. Dynamic learning communities: an alternative to designed intructions. In *Proceedings of Selected Research and Development Presentations, Washington D.C.: Association for Educational Communications and Technology*, pages 800–809.

[18] V. Uskov. A 3rd generation web-based intructional tool for education and lifelong training. *International Journal of Continuing Engineering Education and Lifelong Learning*, 13(1/2):110–131, 2003.

[19] Wikipedia. http://en.wikipedia.org/wiki/Main_Page.

Juegos en Red como Proyecto Docente en Ingeniería Telemática

Sergio Machado*, Roc Messeguer†, Toni Oller*, Angélica Reyes†, David Rincón*, José Yúfera*
*Departamento de Ingeniería Telemática. †Departamento de Arquitectura de Computadores.
Escuela Politécnica Superior de Castelldefels. Universidad Politécnica de Cataluña.
Av. Canal Olímpico, s/n. 08860 Castelldefels (Barcelona)
Teléfono: 93 413 72 06. Fax: 93 413 70 07
E-mail: {smachado, aoller, drincon, yufera}@entel.upc.edu, {messeguer, mreyes}@ac.upc.edu

***Abstract**. The paper describes a cooperative-learning, project-based course currently being offered as one of the specializations of the Telematics Engineering degree at the UPC. The course gathers several topics such as Multimedia, Security, Distributed Applications and Software Engineering, under a common umbrella: a multiplayer network game project. The authors describe the evolution from four isolated, classical subjects to a single, innovative, PBL-based course. The multiplayer game project is described in detail, since it is the linkage point of the course and the main incentive of the students. The game requirements include aspects from the four main topics of the course. The paper ends with a description of the evaluation procedure and the feedback from the students.*

1 Introducción

En este artículo presentamos la asignatura Intensificación en Servicios Telemáticos, un bloque de carácter optativo del tercer curso de la titulación de Ingeniería Técnica de Telecomunicaciones, especialidad en Telemática que se imparte en la Escuela Politécnica Superior de Castelldefels (EPSC) de la Universidad Politécnica de Cataluña (UPC). El bloque consta de 22,5 créditos y está impartido por profesores de dos departamentos: Arquitectura de Computadores e Ingeniería Telemática.

La implantación de este bloque se enmarca en la aplicación de la metodología docente de aprendizaje basado en proyecto (PBL), nueva en el tercer curso de la titulación de la Escuela. Anteriormente a la implantación del bloque existían cuatro asignaturas independientes que con la introducción de la nueva metodología han desaparecido como tales. Debido a ello los profesores han debido adaptar los temarios independientes de tal modo que los objetivos de aprendizaje del bloque fueran vistos como unidades didácticas por los alumnos. Para los alumnos, trabajar con PBL significa realizar una sola matrícula y obtener finalmente una única nota de bloque. Además, se enfrentan a un único proyecto que debe ir desarrollando a lo largo del curso, trabajado en grupo y que pretende aunar esas unidades de didácticas en un solo conjunto de aplicación, si bien está claro que no todas las unidades tendrían reflejo aplicado en el proyecto.

Así pues, el proyecto debía ser concebido de tal modo que integrase la mayor parte de los objetivos docentes planteados para el bloque, evitándose la idea de pensar en dicho proyecto como en una "práctica grande". El proyecto debía ser lo suficientemente amplio como para que los trabajos en grupo estuviesen justificados, y para que el alumno comprendiera las ventajas e inconvenientes del trabajo en paralelo y la posterior integración de resultados, algo que en breve se encontrarán en su futura vida laboral.

Con todo esto, se decidió buscar un proyecto lo suficientemente estimulante como para que el alumno no perdiese el interés durante los cuatro meses de duración del bloque. Para conseguirlo se preparó un proyecto que incluye el diseño y desarrollo de un juego en red multi-jugador y en tiempo real.

Este artículo se centra en presentar las ventajas de un proyecto de tales características para la docencia de aspectos que conciernen a la ingeniería telemática como pueden ser el diseño de protocolos, servicios y aplicaciones, la seguridad, entornos distribuidos y la transmisión de información multimedia.

El artículo se estructura de la siguiente manera: primeramente esbozaremos los objetivos docentes del bloque. Posteriormente estableceremos los requisitos del proyecto del bloque así como su evaluación. Finalmente, presentaremos las impresiones generales de los alumnos recogidas en las encuestas pasadas a los alumnos por parte de la Escuela.

2 Objetivos Docentes

El bloque Intensificación en Servicios Telemáticos se imparte en el tercer curso de la titulación de Ingeniería Técnica de Telecomunicación, especialidad en Telemática, de la Escuela Politécnica Superior de Castelldefels (EPSC). Este bloque unifica en una única asignatura optativa una oferta de cuatro asignaturas optativas anteriormente impartidas en la EPSC. Estas asignaturas eran "Redes y Servicios Multimedia", "Servicios Avanzados de Seguridad", "Transmisión de Información Multimedia" y "Sistemas Distribuidos" (ver Fig. 1). En esta sección primeramente situaremos el contexto inicial del

bloque como cuatro asignaturas separadas, lo que nos servirá para presentar los objetivos docentes del bloque, ya que se debe tener en cuenta que la unificación de contenidos en un único bloque no mermaba los objetivos docentes de cada una de las asignaturas por separado. Seguidamente, se presenta el contexto final del bloque sin diferenciación de asignaturas (sí de contenidos) y con el proyecto aglutinador de los objetivos docentes perseguidos. En [1] y [2] se explican detalladamente las fases que se siguieron desde la situación inicial hasta la final.

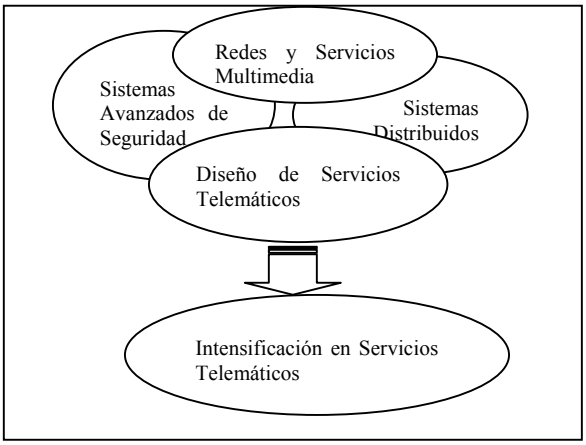


Figura 1- De las asignaturas aisladas al bloque

2.1 Contexto Inicial

"Redes y Servicios Multimedia" tenía como objetivo principal la presentación de todos los aspectos relacionados con la prestación de servicios multimedia, empezando por los algoritmos de codificación, pasando por el formato de la información y su protección, así como por los protocolos de transporte, multicast y calidad de servicio, más allá del "mundo IP", intentando abarcar otras redes de comunicaciones como, por ejemplo, las redes de difusión de televisión terrestre digital, o las redes de acceso vía cable y satélite.

"Sistemas Avanzados de Seguridad" ofrecía conceptos relacionados con la aplicación práctica de la seguridad en el mundo de la ingeniería telemática. Esto incluye trabajar con protocolos de seguridad (SSH, SSL, IPSec), la configuración de cortafuegos avanzados, el funcionamiento de los sistemas detectores de intrusos, o los ataques y vulnerabilidades más comunes en redes de ordenadores. Se daba una visión global de los servicios que requieren de seguridad, como son, el comercio electrónico, la gestión de derechos digitales, etc.

"Sistemas Distribuidos" tenía como objetivo principal aportar a los alumnos los conocimientos básicos sobre sistemas distribuidos, desde fundamentos, pasando por la estructura hasta llegar al diseño. Con más detalle la asignatura presentaba al estudiante los fundamentos, estructuras y diseños de los sistemas de comunicación y colaboración distribuida y abierta, introduciendo al estudiante en el conocimiento del diseño y el funcionamiento de algunas implementaciones y aplicaciones de los sistemas distribuidos utilizados actualmente, permitiéndole experimentar y profundizar en estas implementaciones de los sistemas distribuidos actuales.

"Diseño e Implementación de Servicios Telemáticos" exponía los conocimientos fundamentales del diseño de servicios telemáticos a la vez que se le proporcionaban conceptos de programación en lenguaje Java para la implementación práctica de los diseños. La asignatura constaba de dos partes. En la primera parte se introduce al alumno en los conceptos básicos de la programación orientada a objetos y se le ofrece una presentación del UML (Unified Modeling Language) como técnica de descripción formal para el diseño de servicios telemáticos. La segunda parte proporciona al alumno conceptos de diseño de bases de datos, de la plataforma J2EE (Java 2 Platform, Enterprise Edition) y de programación de red en Java y programación J2EE (Servlets y JSPs), correspondiendo estos dos últimos a los conceptos de programación a los que anteriormente nos referíamos.

2.2 Contexto de Bloque

Una vez comentado el contexto inicial del que partimos para implantar la metodología PBL, comentaremos a continuación el resultado final de los cambios realizados a partir de esta situación.

El objetivo final era el de conseguir un único bloque optativo de 22,5 créditos sin realizar grandes cambios en el plan de estudios en el que se encontraban enmarcadas las cuatro asignaturas mencionadas. El alumno, a la hora de cursar el bloque, debía realizar una única matrícula. Asimismo, los profesores debían evaluar a los estudiantes con una sola nota final, de manera que desapareciesen las fronteras entre asignaturas.

El nexo de unión de todos los contenidos del bloque es el proyecto, en el que aparecerán la mayoría de objetivos docentes del bloque. El proyecto pretendía ser aplicado y tener carácter completo: fase de análisis de requerimientos, de diseño de la propuesta, y de la implementación del diseño y las pruebas de calidad y rendimiento del producto final. De esta forma, los alumnos tendrán que tomar decisiones y probar diferentes tecnologías. Como la metodología PBL indica es un trabajo en grupo, en equipo. La selección de los componentes del grupo no se deja a la libre elección de los alumnos, sino que son los profesores los que deciden la composición de los mismos.

Dos razones, principalmente, avalan la decisión de que sean los profesores quienes formen los grupos. Primeramente, no todos los alumnos matriculados

tienen una carga lectiva igual por lo que los profesores pueden conformar grupos con cargas de matrículas similares. Si bien en el caso ideal un alumno durante el cuatrimestre en el que se imparte el bloque sólo debería estar matriculado de éste y de otra asignatura de la titulación ("Administración de Empresas"), la realidad es que la mayoría de los alumnos suelen ser repetidores de una asignatura del cuatrimestre anterior considerada de alta dificultad por los alumnos, pero que la no asimilación de sus contenidos no interfiere en el correcto seguimiento del bloque (aunque sí en la carga de trabajo que conlleva). Segunda, si uno de los objetivos del PBL consiste en potenciar las capacidades de trabajo en grupo, algo, al parecer, demandado por las empresas, uno no suele decidir cuáles van a ser sus compañeros de trabajo, sino que viene decidido por la propia empresa. Este grupo puede presentar tanto afinidades humanas como no, y aún así, el trabajo debe salir adelante. De todos modos, como se verá, el método de evaluación contempla la posibilidad de que uno o varios elementos del grupo no cumplan con su deber, tratando de corregir esta desviación aunque dejándola en manos de los propios componentes del grupo.

La relación entre los profesores y los grupos se realiza bajo tres roles diferenciados. Los profesores actúan como clientes que contratarán los servicios del grupo para la realización del proyecto que demandan; desde otro punto de vista los profesores tienen el rol de consultor tecnológico para ayudar y resolver problemas a los grupos; por último, cada profesor actúa como supervisor del proyecto para realizar un seguimiento individualizado del trabajo de los alumnos en cada una de las áreas de conocimiento de las que esté encargado de su docencia.

3 El proyecto

En esta sección presentamos y justificamos la elección, como proyecto de bloque, del diseño e implementación de un juego en red multi-jugador con posibilidades multimedia. Veremos cuáles son los requerimientos iniciales que se les da a los alumnos, cómo se va haciendo el seguimiento de la evolución del mismo y en base a qué se obtiene la evaluación final tanto del proyecto, como del bloque.

3.1 Requerimientos del proyecto

Cuando se presenta el título general del proyecto a los alumnos les decimos que éste consiste en el diseño, implementación y pruebas de calidad de un juego en red multijugador, en tiempo real y no por turnos, con capacidades multimedia como chat de voz e imágenes (videoconferencia). Más en detalle, a los alumnos se les proporcionan los siguientes requerimientos:

- **Multijugador**. Se establece un número mínimo de jugadores igual al número de componentes del grupo (5 en nuestro caso). Se debe contemplar la escalabilidad del juego, es decir, cuál sería el número máximo de jugadores que podrían jugar simultáneamente una partida.

- **No por turnos**. Es decir, no vale un juego tipo mus, sino que los jugadores realizan sus acciones concurrentemente.

- **Juegos originales**. En cierta manera se evalúa la originalidad del juego, es decir, que se premia la no adaptación de juegos basados en otros ya existentes.

- **Juego distribuido (Peer-to-Peer)**. Se distingue aquí que el juego puede tener un servidor que sirva para localizar partidas, pero que durante el juego de una partida en particular sólo deben intervenir los jugadores.

- **Acceso heterogéneo a información estadística y otros**. Se debe desarrollar un "portal del juego" que permita acciones como registrase, consultar estadísticas y cualquier otra información que los desarrolladores consideren relevante. Este acceso debe poderse realizar desde terminales heterogéneos: un PC, un teléfono móvil, PDA, etc. Estos accesos deben ser seguros.

- **Autenticación y no-repudio de los datos.** Al final de la partida, que recordemos se juega de modo distribuido, se debe actualizar información como, por ejemplo, los puntos obtenidos por cada uno de los jugadores. Se debe desarrollar algún mecanismo que asegure la autenticidad y el no repudio de esa información.

- **Integridad de la descarga.** Las descargas del ejecutable del juego deben incorporar algún mecanismo que sirva para verificar que la descarga se ha realizado correctamente.

- **Seguridad de datos almacenados.** Se debe crear una red segura, es decir, que, por ejemplo, la base de datos no sea accesible desde el exterior (esta consideración como mínimo).

- **Estudio de la posibilidad de hacer trampas (*cheats*).** Se debe analizar qué posibilidades tiene un jugador de hacer trampas, por ejemplo, conseguir que aunque le hayan matado no se muera para el resto de jugadores, y buscar posibles soluciones que afectarán a los protocolos y sistemas diseñados.

- **Banda sonora recibida por streaming.** El juego no incluye ningún tipo de ambientación musical (sí sonidos). Ésta deberá proporcionarla algún (o algunos) jugador mediante algún sistema de distribución de streaming.

- **Chat de voz.** Los jugadores pueden comunicarse mediante voz durante el juego.

- **(OPCIONAL) Chat de vídeo y voz.** Los jugadores pueden comunicarse mediante vídeo y voz durante el juego incorporando algún mecanismo que transmita vídeo sólo cuando se detecte que se está hablando.

Queremos enfatizar que estos requerimientos no se dan por escrito, sino que se les comenta oralmente y son ellos los que deben ir apuntándolos. Tratamos así de simular la toma de requerimientos que todo proyecto real incluye.

3.2 Motivaciones del proyecto

Cada uno de los requisitos del proyecto persigue una motivación diferente orientada a evaluar la asimilación de los conceptos de las diferentes unidades didácticas, así como a evaluar la capacidad de enfrentarse a un problema telemático complejo como el que proponemos. Es obligado decir que los gráficos del juego no es algo a lo que demos un gran peso evaluativo, pues ni los alumnos tienen por qué tener conocimientos de programas de diseño gráfico ni es un objetivo del bloque. Está claro que visualmente es más atractivo un juego con un nivel gráfico destacable, por lo que se les invita a buscar en Internet mapas y sprites gráficos de libre distribución para que los usen en sus juegos (naves, soldados, mapas). Hasta el momento, el nivel gráfico de los juegos ha sido bastante más que bueno.

La elección general del diseño e implementación de un juego resulta altamente estimulante para los alumnos. Esta propuesta resulta muy atractiva por el hecho de que casi todos los programadores han querido programar su propio videojuego. Como se verá en la Sección 4, esta elección es muy valorada por los alumnos a la hora de elegir este bloque frente a otras alternativas.

Tal y como se plantean los requisitos básicos del juego, es decir, multi-jugador en tiempo real y distribuido, hace que los alumnos se encuentre ante tres problemas telemáticos básicos. Primero, deben afrontar el diseño y la implementación de un protocolo que permita el intercambio de información entre los usuarios. El hecho de especificar que sea en tiempo real invita a los alumnos, aunque no se les especifica, a huir del uso de TCP como protocolo de transporte y elegir UDP. Esta elección, que tiene ventajas evidentes, tiene sin embargo, implicaciones directas sobre campos del protocolo de la capa de aplicación diseñado, como puede ser, por ejemplo, la secuenciación de los paquetes o que, por ejemplo, deberán afrontar la posible pérdida de paquetes. Esta idea se refuerza pasándoles como lectura el artículo [3], recomendado en la bibliografía del bloque. Al pedir que no sea por turnos, los alumnos deben tratar problemas de sincronismo y de tolerancia a retardos. Se les proporciona la referencia [4], que cubre varios aspectos del diseño de juegos en red, para que analicen las implicaciones de no tener actualizaciones síncronas del estado del juego. Por último, el hecho de que sea distribuido les invita a reflexionar sobre las posibles topologías que deben formar los N jugadores: ¿siguen un modelo cliente servidor? En ese caso, ¿quién hace de servidor?, o ¿hacen todos de servidores y clientes como en un peer-to-peer puro? ¿un híbrido, quizá?.

Otro aspecto destacable es la carga de tráfico generada por el juego y la calidad de servicio que se le demanda a la red. A los alumnos se les pide explícitamente que realicen estudios analíticos de carga tanto en la fase inicial de diseño como en la fase final de pruebas. A partir de esto, también pueden explicitar unos requisitos mínimos de red de acceso necesarios para jugar a su juego.

Dentro de las unidades didácticas que correspondían a "Sistemas Distribuidos" se les explica en profundidad lenguajes de representación, como HTML y WML, que correspondería al requerimiento de que se pueda acceder al portal del juego desde terminales heterogéneos. Además, incluye aspectos de diseño de la plataforma J2EE relacionado con patrones de diseño, como el "Modelo-Vista-Controlador" correspondiente a las unidades didácticas de "Diseño e Implementación de Servicios Telemáticos".

Los requisitos que conciernen a aspectos de seguridad persiguen el diseño de un firewall con zona desmilitarizada y, sobre todo, algo muy importante en el juego derivado del hecho de que éste sea distribuido: partiendo de que los jugadores no son de fiar, cuando se almacenen las puntuaciones en la base de datos, éstas deben ser veraces y no que, por ejemplo, si se eligió una arquitectura donde un jugador actúa como servidor, éste pueda alterar las puntuaciones en su beneficio.

Los requerimientos multimedia enfrentan a los alumnos a los problemas derivados de la transmisión audiovisual en tiempo real (streaming de audio, chat de voz y/o vídeo). Se les indica también que lo importante es jugar al juego, es decir, se les está pidiendo algún tipo de mecanismo de control de congestión que deshabilite, temporalmente o no, la transmisión y recepción de este tipo de información en el caso de que las pérdidas y el retardo del juego superen el rango de tolerancia que ellos mismos hayan establecido. Así mismo, esta información aumenta las exigencias a la red pudiendo afectar a los requisitos mínimos anteriormente expuestos.

3.3 Evaluación

Como se ha dicho, el bloque consta de una parte teórica formada por las unidades didácticas y de la parte más aplicada que es el proyecto. La primera cuestión a tener en cuenta es que el proyecto, dada su envergadura, debe tener un peso muy considerable en la evaluación. Nosotros hemos optado por dar un 40% de peso al proyecto y un 60% a las unidades didácticas. El problema está en que un alumno podría muy bien aprobar el bloque sin haber trabajado prácticamente nada en las unidades didácticas (la parte más teórica). A fin de evitar esto se exige haber aprobado las tres cuartas partes de las unidades didácticas para que cuente la nota de proyecto. Esta restricción además consigue que suspender la parte de unidades didácticas correspondiente a las antiguas asignaturas conlleve el suspenso del bloque, con lo que se evita que algún alumno abandone los contenidos de alguna de las antiguas asignaturas.

En cuanto a la evaluación de las unidades didácticas cada profesor responsable de un área de conocimiento evalúa como considere adecuado, individual o colectivamente, mediante trabajos o exámenes, etc.

La evaluación del proyecto consta de dos pruebas. A mitad del cuatrimestre se les pide una presentación de su propuesta donde discutirán con los profesores el diseño y las elecciones que han tomado respecto a los diferentes aspectos del proyecto. Su peso es del 10% (de la nota final). La idea de esta presentación consiste principalmente en corregir a los alumnos los aspectos que los profesores consideramos que les pueden resultar problemáticos en la realización final del proyecto.

A final del cuatrimestre se les pide una memoria que vale el 15%. En ella deben reflejar sus propuestas, sus soluciones, sus pruebas de calidad y demostrar que han cumplido con los requisitos iniciales del proyecto. El 15% restante se evalúa con la presentación final (un resumen de la memoria) y una demostración de su implementación. De cara a esta última evaluación, los profesores preparan un entorno de red "hostil", pudiendo así comprobar que las afirmaciones de los alumnos, en cuanto a tolerancia de pérdidas y retardos, son correctas.

La nota que se les da a cada uno de los grupos en la evaluación del proyecto es la nota media del grupo, es decir, que luego ellos se reparten la nota de tal modo que la media de la nota de los alumnos debe ser igual a la nota que los profesores dan. Con esto tratamos de que sea el grupo mismo quien valore la aportación individual de cada uno de los miembros, cosa que puede quedar fuera de las posibilidades evaluativos de los profesores. Grupos que se han encontrado con algún miembro (o miembros) que no han producido lo que se esperaba de ellos, pueden castigar a esos elementos bajándoles la nota y subiéndosela a ellos. Del mismo modo, si el grupo reconoce que ha habido alguien clave en la consecución del resultado final pueden premiarle bajándose ellos un poco la nota y dando puntos a ese miembro. Y en caso de que las contribuciones hayan sido igualitarias, todos se quedan con la misma nota.

Este último punto queda bien claro en la presentación del bloque que se les hace a los alumnos, de modo que se impulsa desde el principio el que se repartan bien las tareas entre los miembros de un grupo.

4 Valoración de los Estudiantes

Durante los diferentes cuatrimestres en los que se ha realizado el proyecto del bloque optativo, los alumnos han rellenado diferentes encuestas. A nivel de centros, por un lado están las encuestas de la Universidad Politécnica de Cataluña y por otro las propias de la Escuela Politécnica Superior de Castelldefels.

En las primeras, tipo test, los alumnos contestan a una serie de preguntas comunes a todas las asignaturas de la universidad. Por el contrario, en las de la escuela, los alumnos tienen libertad para mostrar sus impresiones libremente, guiados únicamente por dos opciones: las cosas favorables y las desfavorables de la asignatura que están evaluando. Además de esto, se les pide que indiquen la carga de trabajo que dicha asignatura supone para ellos, incluyendo las horas lectivas.

Además de estos dos casos, los profesores del bloque optativo de Intensificación en Servicios Telemáticos han pasado a los alumnos una adaptación del cuestionario SEEQ (Student Experience of Education Questionnaire) en el que se hace hincapié en el trabajo en grupo [5].

En las encuestas se observa una buena recepción por parte de los alumnos hacia los temarios del bloque. A parte de esto, a continuación destacaremos los resultados más relevantes que se pueden obtener de las respuestas a dichas encuestas y que pueden ser de interés tanto en la aplicación de técnicas PBL como para valorar el bloque como unidad cuyo motor es el proyecto de diseño de juegos en red.

- La carga del bloque, en horas semanales de dedicación incluyendo las lectivas, es considerada por la mayoría de los alumnos como elevada para el número de créditos que tiene el bloque asignados;

- Como aspectos más positivos los estudiantes han valorado el trabajo en grupo, la realización de informes y presentaciones, y el grado de libertad con el que se encuentran a la hora de tomar decisiones para realizar un proyecto con requisitos fijados, pero

desarrollo poco guiado a priori. También consideran importante la libre disponibilidad del laboratorio en el que pueden implementar las diferentes partes del proyecto. Por otra parte, han valorado positivamente la coordinación entre los diferentes profesores y unidades del bloque optativo. En general, consideran que la cantidad, calidad y utilidad de los conocimientos adquiridos es elevada.

- Como aspectos más negativos: la inclusión en el bloque de unidades didácticas (o parte de ellas) que no se ven directamente reflejadas en el proyecto, pero que se consideran importantes para conseguir los objetivos del mismo. Cabe destacar también la sensación que tienen sobre el poco tiempo que pueden dedicarle al proyecto, por culpa de la carga que supone la evaluación (mediante trabajos, prácticas o controles) de las unidades didácticas que componen el bloque.

Consideramos que este último punto se debe, principalmente, a no haber guiado un poco más a los alumnos en el desarrollo del proyecto durante la primera fase del mismo: el diseño. Los alumnos tienden a centrarse más y aplicar más esfuerzos en la fase de implementación, no comprendiendo la importancia de la fase previa. Además, un error en el diseño del juego puede suponer un cambio que debe realizarse en un tiempo más limitado. En estos momentos, los profesores del bloque de Intensificación en Servicios Telemáticos intentan controlar más la evolución del proyecto desde el primer día.

De cara al profesor este control conlleva una caga adicional de trabajo docente que no es despreciable, pero sí asumible, teniendo en cuenta que en estos momentos hay cinco profesores asignados (no en exclusividad, ya que también imparten docencia en otras asignaturas), lo que permite repartir la carga. Por otro lado los profesores se benefician del período de aproximadamente 15 días al final del cuatrimestre, en el que se dan por finalizadas las sesiones teóricas y los estudiantes se dedican por completo a la programación del proyecto.

5 Conclusiones

En este artículo hemos presentado el bloque "Intensificación en Servicios Telemáticos" que se imparte en el tercer curso de la titulación de Ingeniería Técnica de Telecomunicación, especialidad en Telemática. El bloque se imparte según los principios de la metodología PBL en la que básicamente el aprendizaje se estructura entorno a un proyecto.

Hemos presentado nuestra elección como proyecto el diseño e implementación de un juego en red multijugador, distribuido y en tiempo real. A través de los comentarios recogidos en diversas encuestas realizadas a los alumnos del bloque se ha demostrado que esta elección resulta altamente motivadora y se adecua, perfectamente, tanto a los objetivos planteados por la enseñanza basada en PBL, como a que los alumnos asimilen los objetivos docentes que se imparten.

No disponemos de comparaciones de rendimiento entre el método tradicional y el basado en proyecto, pero por la experiencia de los últimos cuatrimestres estamos convencidos (y, lo que es más importante, nuestros estudiantes también lo están) de que el bloque proporciona a los alumnos una formación teórica y práctica que va mucho más allá de la que se que puede conseguir con un método de enseñanza tradicional.

Referencias

[1] S.Machado, R.Messeguer, A.Oller, M.A.Reyes, D.Rincón y J.M.Yúfera, "Recomendaciones para la Implantación del PBL en Créditos Optativos Basadas en la Experiencia en la EPSC", XI Jornadas de la Enseñanza Universitaria de la Informática (Jenui), Julio, 2005.

[2] S.Machado, R.Messeguer y J.M.Yúfera, "Juegos en Red: una Experiencia de Enseñanza Basada en Proyecto en la EPSC", XII Congreso Universitario de Innovación Educativa en las Enseñanzas Técnicas (CUIEET), Julio, 2004.

[3] P.Lincroft, "The Internet Sucks: Or, What I Learned Coding X-Wing vs. TIE Fighter". Disponible en: http://www.gamasutra.com/features/19990903/lincroft_01.htm. Septiembre, 1999.

[4] J.Smed, T.Kaukoranta and H.Hakonen, "A Review on Networking and Multiplayer Computer Games", Technical Report 454, Turku Centre for Computer Science, 2002.

[5] SEEQ at Curtin University of Technology, http://lsn.curtin.edu.au/seeq/.

OpenSimMPLS: Herramienta para la Innovación Docente e Investigación en Redes y Comunicaciones[Ú]

F. Rodríguez-Pérez[1], M. Domínguez-Dorado[1], J. L. González-Sánchez[1], J. L. Marzo Lázaro[2], A. Gazo-Cervero[1]
[1]Dpto. de Informática. Universidad de Extremadura, Avda. de la Universidad s/n, CP-10071 Cáceres
[2]Institut d'Informàtica i Aplicacions, Universitat de Girona, Girona
fjrodri@unex.es, ingeniero@manolodominguez.com, jlgs@ex.es, marzo@eia.udg.es, agazo@unex.es

Abstract. MPLS (Multiprotocol Label Switching) provides interesting mechanisms to integrate network technologies like ATM and IP with Quality of Service. It is a last generation technology that presents great interest for teaching at the university. In this project, a simulator, as a didactic resource for MPLS teaching innovation, is presented. It allows to the student to configurate, interact and analyze the operation of a MPLS domain in a simple and efficient way. On the other hand, due to its free software license, it can also be used as a protocol engineering platform for creating of analysis and results validation tools by researchers who work in MPLS related projects.

1 Introducción

MPLS (Multiprotocol Label Switching) [1] es una tecnología orientada a conexión que surge para orientar los problemas que plantean las redes actuales en cuanto a velocidad, escalabilidad e ingeniería de tráfico [2]. Al mismo tiempo ofrece Calidad de Servicio (QoS) extremo a extremo, mediante la diferenciación de flujos y reserva de recursos. Por otro lado, elimina el problema de la gestión de los diferentes planos de control que tienen lugar en redes IP/ATM, proporcionando mecanismos para conseguir la convergencia entre ambas tecnologías.

MPLS actúa como nexo entre los protocolos de red y el correspondiente protocolo de nivel de enlace. Para ello, en la estructura de una trama, la cabecera MPLS se situará después de la cabecera de nivel de red y antes de la cabecera de nivel de enlace [3]. De hecho, el reenvío de paquetes MPLS está basado en etiquetas y no en el análisis de los datos que encapsulan (protocolos de nivel de red).

Es una tecnología multiprotocolo, como su propio nombre indica; admite cualquier protocolo de red, pero al mismo tiempo permite cualquier tecnología de capas inferiores (enlace o físico). De esta forma, se ha proporcionado un atractivo mecanismo para aprovechar la infraestructura actualmente desplegada en ámbitos troncales, facilitando así la migración de tecnologías [4]. Los esfuerzos realizados desde hace años para desarrollar mecanismos innovadores que den soporte a IP sobre ATM no se han perdido, ya que la mayoría de las técnicas desarrolladas serán válidas para disponer de IP sobre MPLS y MPLS sobre ATM.

En el presente artículo se presenta un simulador MPLS al que hemos denominado OpenSimMPLS. Es una herramienta funcional y visual que puede utilizarse en la docencia de asignaturas de redes y/o comunicaciones [5]. Contempla los aspectos fundamentales de funcionamiento y configuración de un dominio MPLS [6]; al mismo tiempo ha sido mejorado al incluir compatibilidad con dominios que soporten Garantía de Servicio (GoS) [7]. Un dominio MPLS con capacidad GoS puede entenderse como un entorno capaz de llevar a cabo recuperaciones locales de paquetes descartados junto con la posibilidad de recomponer localmente LSPs (Label Switched Paths) [8]. Éste es el ámbito actual de nuestra investigación y OpenSimMPLS está siendo empleado para validar los resultados obtenidos. En [9] y [10] podemos encontrar otros simuladores MPLS que permiten el diseño y configuración de los componentes de un dominio, así como la simulación y análisis estadístico de los resultados, todo ello desde un punto de vista docente. En el caso de OpenSimMPLS, también serán usuarios del simulador los investigadores de proyectos MPLS con necesidad de comprobar la bondad de sus resultados. Para flexibilizar el futuro empleo del simulador por diferentes investigadores, el software es multiplataforma y está liberado bajo licencia GPL v2.0 de *Free Software Foundation.*

En la siguiente sección se hace una breve descripción del entorno visual del simulador, así como de algunos aspectos funcionales. En el tercer apartado se hace hincapié en la importancia de OpenSimMPLS en entornos docentes o para la validación de resultados de investigación. En el cuarto apartado se comentan algunos detalles generales relativos a su programación. Finalmente, el artículo concluye indicando las contribuciones del simulador.

2 Aplicaciones docentes e investigadoras de OpenSimMPLS

OpenSimMPLS es una herramienta multiplataforma, cuyo objetivo fundamental es el de servir al profesorado universitario como recurso didáctico para innovar en la docencia y análisis del funcionamiento de redes MPLS. El simulador dispone de una interfaz gráfica que permite un

entorno de usuario simple. La programación de cada uno de los elementos que componen la aplicación está orientada a objetos; además genera procesos que funcionan de manera concurrente mediante hilos independientes, lo cual permite el estudio de los distintos eventos de una forma más flexible. En cuanto a su funcionamiento, el simulador dispone de tres niveles de operación: el primero trata el diseño y configuración de la topología de un dominio MPLS; el segundo considera la exploración visual de los diferentes eventos que van sucediéndose durante la simulación; y el tercero permite la evaluación de las prestaciones del dominio MPLS diseñado.

El alumno de asignaturas de redes y comunicaciones refuerza su aprendizaje gracias a ejemplos prácticos, ya que el simulador ofrece resultados sobre el comportamiento de la red al introducir servicios particulares como, por ejemplo, tráfico multimedia. De esta forma puede apreciarse cómo se priorizan unos flujos con respecto a otros, permitiendo contrastar resultados gracias al sistema de reconfiguración de los dispositivos previamente definidos en el dominio. También se puede modificar la topología, incorporando otros elementos para identificar los parámetros físicos de diseño de la red que garantizan una mayor QoS. De esta forma el alumno puede realizar propuestas para la mejora de supuestos de redes MPLS, con el objetivo de establecer los criterios mínimos de calidad de servicio. Así, el estudiante puede analizar y diseñar la integración de servicios de banda ancha (multimedia) en redes MPLS, para conocer el comportamiento de éstas, la calidad del servicio ofrecida y dónde se producen los efectos más perniciosos sobre el tráfico, con el objetivo de aprender a evitarlos en lo posible.

En cualquier simulación MPLS, el alumno resuelve problemas y situaciones, aprende procedimientos, llega a entender las diferentes características de los eventos, aprende cómo controlarlos y qué acciones realizar en situaciones particulares. Se puede emplear OpenSimMPLS de forma que el estudiante se trace hipótesis basadas en su experiencia y conocimientos teóricos acumulados, a modo de síntesis o repaso de lo que ya ha estudiado. Tiene la posibilidad de poner en práctica sus ideas, obtiene información de retorno inherente a las respuestas del simulador, las cuales debe descifrar para saber qué ocurre en el interior del dominio y determinar cuál es la norma o principios que rigen su funcionamiento. Este proceso experimental y analítico es el que ayuda al estudiante a desarrollar sus propias estrategias de pensamiento. Un ejemplo de uso del simulador en el aula podría consistir en proponer el diseño de dominios MPLS que presenten situaciones conflictivas de congestión. Dicha problemática deberá ser detectada por el alumno durante la simulación del sistema, así como en el Área de Análisis de resultados (estadísticas de paquetes entrantes, descartados, salientes, etc.). El alumno también deberá plantear medidas para solventar la problemática localizada: modificación de la topología, cambios en las características de los tráficos, empleo de técnicas de priorización de flujos, etc. De igual forma, podrá comprobar cómo las variaciones introducidas contribuyen a mejorar el rendimiento final del sistema.

OpenSimMPLS permite al usuario variar múltiples parámetros de configuración para intervenir directamente en las operaciones simuladas en cualquier instante, ofreciendo una visión interactiva de los diferentes eventos. De esta forma, la característica más importante de este simulador interactivo es permitir al estudiante ser un agente activo durante la simulación, profundizar en los procesos simulados de funcionamiento MPLS, para luego saber interpretar y manipular los resultados que se obtengan durante la ejecución. El alumno, además de participar en la situación, debe ser también capaz de procesar la información que el simulador le proporciona, propiciando el conocimiento de tipo experimental y el aprendizaje por descubrimiento.

Debido a la dificultad para la implementación de todo este tipo de pruebas sobre redes MPLS reales, OpenSimMPLS es una solución para enfrentar al estudiante de redes con los diferentes casos de funcionamiento que plantearía un posible dominio MPLS. Así, el docente puede investigar, planificar y dimensionar los recursos de una red MPLS sin necesidad de correr riesgos al modificar las configuraciones de dispositivos reales. Por otra parte, y como objetivo colateral, los resultados analíticos obtenidos pueden ser de utilidad al alumno para desarrollar nuevas metodologías de diseño de cara a su futuro trabajo como planificador de redes de última generación. Para los estudiantes y futuros profesionales es, por tanto, muy útil la adquisición de una disciplina que primero conlleva la simulación de la topología de red, la obtención de resultados y los sucesivos refinados hacia la arquitectura final antes de su implantación, la cual implica una racionalización de la infraestructura MPLS de telecomunicaciones necesaria, así como una estimación del coste o inversión que ésta puede implicar.

Por tanto, el empleo del simulador puede ser un aspecto crucial para la comprensión clara de los conceptos de comunicaciones en MPLS. Refuerza los conceptos teóricos de la tecnología y ofrece al estudiante una motivación a la hora de comprender la interacción entre los diferentes componentes que forman un escenario. La simulación de un dominio MPLS es un mecanismo para orientar la comprensión del funcionamiento del protocolo, así como de las diferentes cuestiones de diseño y sus repercusiones en el rendimiento.

En resumen, como apoyo a la docencia sobre redes MPLS, el empleo del simulador presenta diversas ventajas:

- Con OpenSimMPLS se puede modificar la configuración de los distintos componentes, para

luego analizar las consecuencias de dichos cambios. Sobre una red MPLS real no siempre estará permitido realizar esos cambios de configuración.

- La simulación permitirá la obtención de estadísticas detalladas, las cuales se pueden utilizar para comprobar posibles compromisos de QoS sobre tráficos particulares.

- El empleo del simulador supondrá siempre una solución docente y de validación del aprendizaje más económica que la implantación de un dominio MPLS real.

En nuestro caso, OpenSimMPLS se utiliza para innovar en la docencia de asignaturas universitarias de redes de ordenadores, tales como *Comunicaciones en Banda Ancha* o también *Planificación, Especificación, Diseño y Evaluación de Redes*, ambas impartidas en el segundo ciclo de Ingeniería Informática, en la Universidad de Extremadura. También se considera un recurso idóneo para impartir cursos de doctorado o líneas de investigación sobre redes avanzadas.

2.1 Validación de resultados

OpenSimMPLS ofrece también la posibilidad de ser usado como herramienta de validación de resultados en proyectos de investigación. Por tanto, el otro gran colectivo de usuarios estaría formado por los investigadores que deseen disponer de una herramienta de simulación potente para avanzar en la generación de nuevas técnicas MPLS, basadas en propuestas diferentes a las actuales, con las consiguientes medidas de prestaciones de estos nuevos conceptos en una red. El simulador se ha desarrollado como software libre y así se facilita su empleo como plataforma para la ingeniería de protocolos: el investigador de proyectos relacionados con MPLS puede añadir nuevas clases o módulos al proyecto en función de sus necesidades. De esta forma puede obtener una extensión del simulador original para dar cabida al análisis de los resultados específicos de sus investigaciones.

Para nuestra investigación, el simulador ha sido ampliado de forma que ofrezca compatibilidad con escenarios capaces de ofrecer Garantía de Servicio (GoS). Podemos entender como GoS a la posibilidad de ofrecer a un cierto flujo de tráfico la seguridad de que en condiciones normales será tratado preferentemente con respecto al resto de flujos y que, en caso de la existencia de problemas (pérdidas de paquetes y caídas de enlaces), la arquitectura pondrá los medios necesarios para que en cualquier caso el flujo que requiera GoS sea favorecido, tanto más cuanta más GoS haya sido especificada para ese flujo y siempre en función de las posibilidades de los nodos que atraviese. En la figura 1 se presenta, a modo de ejemplo, una topología en la que cuatro nodos emisores generan flujos de diferentes niveles de GoS. Todos los LSP creados atraviesan un LSR (Label Switch Router) central, con el signo de admiración, el cual sufrirá, por tanto, una elevada congestión. Como consecuencia de esto, se producirá el descarte de paquetes en dicho nodo.

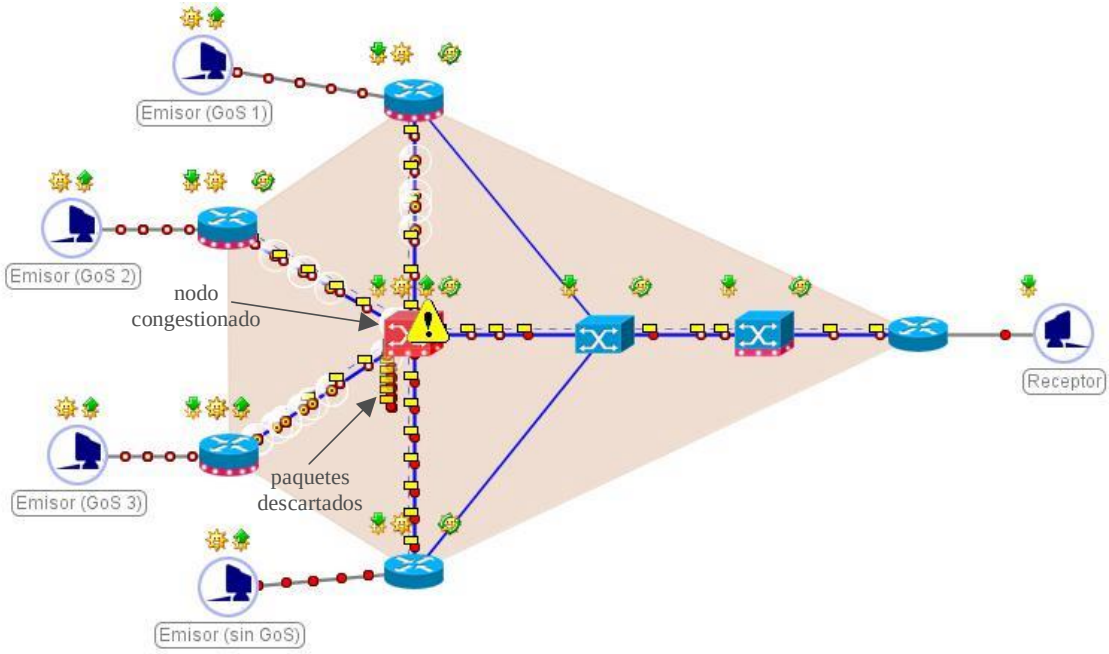


Figura 1: Escenario en el que coexisten flujos de diferentes niveles de GoS

En la gráfica de la figura 2 se muestra que los flujos más prioritarios (con mayor nivel de GoS) tendrán un menor número de paquetes descartados. Es decir, estos flujos se priorizan con respecto a los de menor GoS. De esta forma, el simulador soporta el trabajo con varios niveles de GoS. Por tanto, cada paquete puede ser marcado con estos atributos desde el nodo inicial hasta el nodo final. Cada uno de estos niveles debe entenderse como un grado de probabilidad de que un paquete se pueda localizar en cualquiera de los nodos activos por los que ha pasado (LERA, LSRA). Así se evita la necesidad de retransmisiones extremo a extremo. Por otro lado, se ha considerado también la posibilidad de recuperación de LSPs rotos de forma local adoptando una solución en un entorno mucho más local. Para validar los resultados de nuestro trabajo, hemos utilizado el simulador para comparar de manera estadística el rendimiento que ofrece el simulador para una red MPLS con los resultados de un dominio con soporte de flujos privilegiados GoS.

En resumen, además de los docentes y estudiantes universitarios, existen otros beneficiarios muy directos, entre los que se puede citar a los investigadores que centran su trabajo en estos temas, tanto en universidades como en centros de investigación o escuelas técnicas.

3 Simulador OpenSimMPLS

OpenSimMPLS está constituido por una aplicación *jar* auto-contenida. Su instalación, por tanto, no requiere de ningún paso significativo, y simplemente se debe invocar su ejecución a través de la Máquina Virtual Java de SUN que debe haber sido previamente instalada.

La principal característica del entorno de trabajo del simulador se basa en su simplicidad. Se divide en tres partes: *área de trabajo*, *menú principal* y *ventanas de escenarios*. El área de trabajo es el entorno principal, dentro del cual se desarrollará la simulación de los diferentes escenarios MPLS. El menú principal está situado en la parte superior izquierda, de forma similar al de cualquier otra aplicación, englobando las opciones relacionadas con la gestión de ficheros (crear, almacenar y recuperar escenarios de disco), visualización de ventanas y ayuda.

Figura 2: Paquetes descartados en el LSR Activo

Por último, las ventanas de escenarios permitirán el diseño y análisis de escenarios MPLS particulares. Su estructura se divide en varias pestañas (Fig. 3), las cuales se irán describiendo en los siguientes apartados.

3.1 Área de Diseño de Topologías

La primera pestaña engloba el *Área de Diseño*, en la que se establecerán los parámetros relacionados con la topología y configuración del dominio a simular. La barra de herramientas incorpora diversos iconos que representan los elementos que se pueden insertar en un dominio de OpenSimMPLS (Fig. 3).

El primer icono hace referencia al *Emisor*, que es el tipo de nodo encargado de generar tráfico de red en el simulador. El segundo icono referencia al *Receptor*, el cual actuará como sumidero del flujo generado por un emisor. El tercero representa los *LER* (Label Edge Router), encargados de etiquetar paquetes IP o MPLS, clasificarlos, establecer un camino hacia el destino a través del dominio MPLS y, finalmente, permitir la entrada del paquete etiquetado al dominio MPLS. El cuarto icono es el *LERA* (Label Edge Router Activo), que realiza la misma tarea que el *LER*, pero además se encarga de analizar la cabecera IP para saber si los paquetes tienen requerimientos de garantía de servicio (GoS) y si es así, codificar esos requisitos en la cabecera MPLS [7].

Un flujo IP marcado con GoS sólo puede conservar esos atributos de GoS dentro del dominio MPLS si accede a él a través de un nodo LERA. El siguiente icono representa al LSR, encargado de conmutar tráfico MPLS en el interior del dominio. Es un componente muy rápido, pues sólo observa la etiqueta puesta sobre el paquete por el LER/LERA de entrada al dominio MPLS. Un nodo LSR nunca puede hacer de nodo de entrada al dominio MPLS pues no tiene capacidad para ello. El sexto icono hace referencia a los LSRA (Label Switch Router Activo), que serán los encargados de conmutar tráfico MPLS en el interior del dominio. Además, el LSRA es el componente con capacidad de recuperación local de paquetes y de reestructuración de caminos (LSP) en un entorno local. También tendrá capacidad de almacenar paquetes de forma temporal, para así satisfacer las posibles solicitudes de retransmisión local de otro LSRA del dominio. El último icono representa al *Enlace*, que es el elemento que une dos nodos cualesquiera de la red. Todo escenario de simulación debe tener sus componentes conectados mediante enlaces, por los que fluye el tráfico.

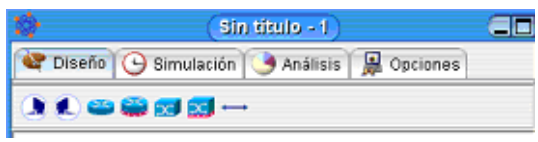


Figura 3: Detalle de la ventana de escenarios de OpenSimMPLS

3.2 Área de Simulación de Escenarios

Debe pasarse al *Área de Simulación* cuando se ha finalizado la creación de la topología del escenario. En este entorno puede analizarse visualmente el comportamiento de dicha topología una vez comiencen a generarse tráficos, saturaciones, caídas de enlaces, etc (Fig. 4). La topología de simulación que podremos ver será la correspondiente al escenario que se haya diseñado en el área de diseño. El área de simulación presenta una estructura similar a la de diseño. La diferencia estriba en que en el lugar donde aparecían los elementos a insertar en el escenario, ahora aparecen unos iconos para controlar el funcionamiento de la simulación.

Si se ha terminado de diseñar y configurar la topología en el área de diseño, se puede poner en funcionamiento la simulación. Esto se hace mediante un clic en el primer icono, que simula un engranaje. Cuando la simulación está en funcionamiento, una barra de progreso indica en todo momento el porcentaje de la simulación en curso. También existe un contador que muestra el número de nanosegundos consumidos en la simulación. Por otro lado, también es posible ralentizar la simulación, lo cual es muy útil para observar con detenimiento los sucesos que van ocurriendo, sin necesidad de detener y reanudar periódicamente la simulación. De esta tarea se encarga el deslizador que se encuentra en la barra de herramientas (Fig. 5).

Toda la simulación visual que se puede observar en el área de simulación en tiempo real no es sino la representación gráfica de los valores internos generados por los elementos que componen el escenario. En la mayoría de las ocasiones, la simulación visual, junto con las estadísticas generadas por los nodos que estén configurados para ello, es suficiente para comprender los diferentes eventos ocurridos en la simulación. Sin embargo, hay ocasiones en que es necesario tener la posibilidad de acceder a una interpretación numérica de alguna situación compleja.

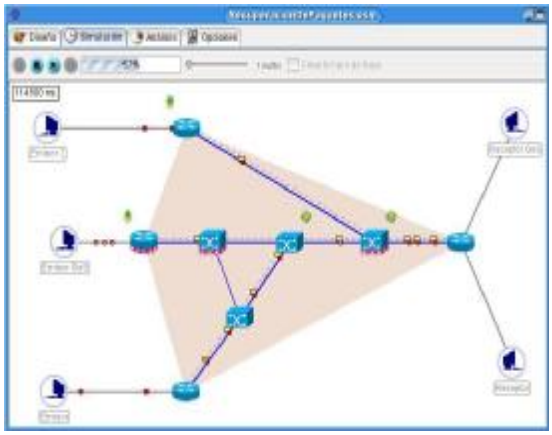


Figura 4: Área de Simulación, en ejecución de un escenario de ejemplo

Figura 5: Detalle de la barra de herramientas del Área de Simulación

Para ello, es posible generar un fichero de traza donde se almacenen, en formato de texto plano, todos los eventos que han tenido lugar durante la simulación: qué componente se ha visto afectado, en qué instante, consecuencias, etc. De esta forma se proporciona un método funcional para la revisión a posteriori de la simulación. Si se desea generar este fichero de traza, se debe hacer clic con el botón principal del ratón sobre el recuadro de selección llamado “Crear fichero de traza”, de la barra de herramientas.

Durante la simulación, los diferentes elementos del escenario podrán ir modificando su aspecto visual a medida que se avanza en el tiempo. Por ejemplo, los nodos LER y LSR modificarán su color en función del nivel de congestión que sufran.

El cambio de una apariencia a otra se realiza de forma automática, a medida que los paquetes se van acumulando en el búfer del nodo. Los paquetes permitirán conocer qué tipos de flujo (clasificados según su prioridad) se dan en el escenario. También informan sobre la cantidad y tipos de tráficos que se mueven por la red, cuándo y cómo se produce la señalización, caminos por los que circulan, velocidad a la que se mueven, etc.

Por otro lado, además de circular por la red que se esté simulando, los paquetes podrán ser descartados en nodos que sufran un elevado nivel de congestión. En ese caso los paquetes aparecerán, visualmente, cayendo de dicho nodo (Fig 6).

Los diferentes aspectos comentados sobre la representación de paquetes pertenecientes a diferentes tipos de tráfico, así como del flujo de los mismos, se puede consultar durante la simulación, gracias a la leyenda que se muestra (opcionalmente) en la esquina inferior derecha del entorno (Fig 7).

Figura 6: Descarte de paquetes pertenecientes a diferentes tipos de tráfico

Figura 7: Leyenda informativa sobre tipos de paquetes, flujos y LSP

Hasta ahora, se han analizado algunas de las señales visuales que se deben interpretar durante la simulación para comprender los sucesos que tengan lugar. Sin embargo, la simulación es un entorno interactivo, podrán llevarse a cabo numerosas acciones durante el funcionamiento de la misma. Por ejemplo, además de mostrarse el nivel de congestión de un nodo particular, puede provocarse una congestión haciendo clic con el botón principal del ratón.

A partir de ese momento, el nodo experimentará una elevada saturación de paquetes (Fig 8). Lo habitual es que en un corto periodo de tiempo el nodo comenzará a descartar paquetes si sigue recibiendo tráfico entrante. Esta función es muy práctica para provocar pérdidas y recuperaciones de paquetes sin tener que esperar la congestión del nodo.

En condiciones reales, un enlace está sujeto a la posibilidad de averías. Obras, descargas eléctricas, fallos humanos, etc., pueden hacer que un enlace falle y el tráfico se pierda. Sin embargo, el exceso de tráfico no provoca el fallo del enlace. OpenSimMPLS permite simular este hecho, admite que un enlace pueda caer en un momento dado, pero al igual que en la realidad, tampoco es algo que ocurra como evolución de la simulación sino que se ha de provocar manualmente. Podremos simular la caída de un enlace durante una simulación mediante un clic de ratón sobre el mismo (Fig. 9). El enlace cambiará su apariencia, mostrándose como una línea roja discontinua y provocando que todos los paquetes circulantes sean descartados. De esta forma podrán simularse situaciones en que el dominio deba recuperarse de fallos de enlace.

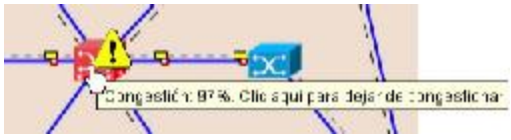


Figura 8: Congestión/descongestión artificial de nodos

Figura 9: Rotura manual de enlace

3.3 Área de Análisis de resultados

Puede pasarse a trabajar al Área de Análisis cuando se han configurado en la topología algunos elementos para que generen estadísticas y se desee observar las gráficas que generan (o que ya han sido generadas, si la simulación ha concluido). Este área se divide en dos partes: una barra de herramientas para el análisis y una zona mayor donde se mostrarán las estadísticas solicitadas para los componentes particulares (Fig 10). Si la simulación está en curso, las gráficas tendrán un comportamiento dinámico, variando según va evolucionando la simulación. Si ya ha finalizado, las gráficas mostrarán los resultados definitivos.

Por otro lado, las gráficas generadas por OpenSimMPLS para cada elemento no son imágenes estáticas, sino que actúan como objetos interactivos. Se puede obtener un menú emergente con opciones sobre cada una de las gráficas, simplemente haciendo clic sobre ellas con el botón secundario del ratón (Fig. 10). De esta forma se tendrá acceso a diversas funciones, como almacenar la imagen en disco, hacer zoom o imprimir una gráfica de interés, entre otras.

4 Detalles de implementación

Una de las ventajas de OpenSimMPLS es su portabilidad, ya que funciona de forma independiente a la arquitectura o sistema operativo del ordenador en el que se ejecute. Para ello se ha empleado el lenguaje Java. Éste también ha permitido la implementación del simulador como una aplicación multiproceso, mediante la programación de hilos.

Figura 10: Área de Análisis, con detalle de menú emergente sobre una gráfica

Asimismo, Java es un lenguaje orientado a objetos; la clase principal del sistema, denominada *openSimMPLS*, inicia la ejecución del simulador. El método *main()*, que se encuentra en esta clase, crea un objeto de tipo *TDispensadorDeImagenes* que será el encargado de cargar todas las imágenes necesarias en la aplicación y que posteriormente será pasado como parámetro en el constructor de cualquier elemento referente a la interfaz. Posteriormente se crea un objeto de tipo *JSimulador* que es la interfaz principal de la aplicación. A partir de este momento la ejecución del simulador dejará de ser secuencial y en su lugar atenderá a los eventos generados por el usuario en la interfaz: órdenes de ratón, selección de opciones de menú, etc.

Durante la simulación, un componente *reloj* enviará avisos a los elementos de la topología (enlaces y nodos) en forma de eventos de temporización o tics (Fig. 11). El reloj es un elemento que se configura con dos valores: por un lado, la duración total de la simulación completa (número de tics); por otro lado, la duración de cada tic. El reloj, que se ejecuta en un hilo propio, avanzará desde cero hasta el número máximo de tics definido para la simulación completa.

Cada tic será enviado a todos los elementos de la topología, de manera que al llegar al máximo de duración de la simulación, el hilo se detendrá y se finalizará la simulación.

Cuando los diferentes elementos de la topología reciben un tic, también reciben la duración en nanosegundos del mismo. En este momento, cada componente activa su propio hilo de ejecución, generándose por tanto concurrencia. Cada hilo realizará su función correspondiente, en función del tipo de dispositivo; por ejemplo, conmutar, transportar paquetes, recibir tráfico, etc. El hilo de ejecución de cada elemento se detendrá cuando la duración del tic recibido se agote.

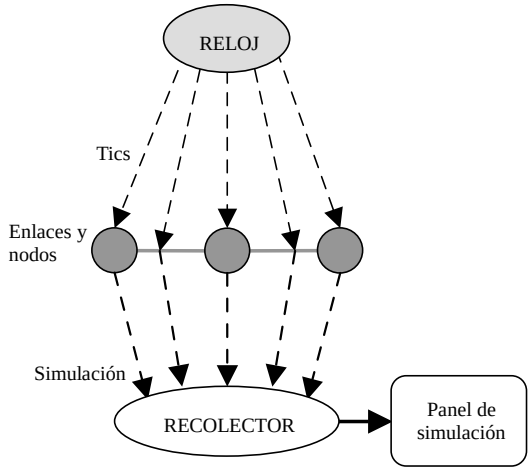


Figura 11: Esquema general de funcionamiento de la simulación de un escenario

Cuando todos los elementos han agotado su tic, el reloj lo detecta y genera el siguiente tic y la operación se vuelve a repetir. La clase que implementa el reloj del sistema se denomina *Treloj* e incorpora una lista interna de todos los elementos a los que debe enviar los eventos de temporización.

Durante el tiempo que el hilo específico de cada elemento está en funcionamiento, se dan multitud de sucesos que deben ser recogidos. Una vez representados, darán lugar a la visualización en pantalla de cada suceso. Este proceso de adquisición lo realiza el *recolector* global del escenario. Todos los elementos de la topología le notificarán la labor que están llevando a cabo durante su tiempo de funcionamiento. Mediante la reiteración de este proceso se consigue una simulación muy fluida, aunque la generación de tics sea un proceso discontinuo. Por tanto, los diferentes elementos de una topología (*TNodoEmisor*, *TNodoReceptor*, *TNodoLER*, *TNodoLSR*, *TNodoLERA*, *TNodoLSRA*, *TenlaceInterno*, *TenlaceExterno*), encapsularán métodos para llevar a cabo estas tareas.

La clase que implementa el recolector de eventos de simulación se denomina *TrecolectorSimulacion*. Implementa el método *capturarEventoSimulacion()*, que permite que los elementos de la topología puedan enviarle los eventos de simulación que van generando durante los tiempos en que sus hilos están en ejecución. Sin embargo, el recolector no muestra los sucesos, sólo los recoge. Para la visualización se han de utilizar los servicios de un componente gráfico que vaya representando en pantalla los diferentes eventos, el cual recibe el nombre de *Panel de Simulación*. De esta forma se consigue aislar las tareas de generación y recopilación de sucesos de las de representación visual de los mismos.

La pantalla de simulación está implementada en la clase *JPanelSimulación*, que realiza todas las operaciones de refresco de pantalla, simulación visual, etc.; es decir, interpreta todos los eventos que le llegan al recolector y los muestra en pantalla de una forma accesible para el usuario. Encapsula *ponerEvento()*, que es el método usado por el recolector para enviarle los eventos que ha adquirido.

4.1 Topología del escenario

La topología es un objeto que almacena todos los elementos del escenario y que se encarga de conectar o interconectar enlaces con nodos y de establecer las asociaciones entre los elementos y el reloj o entre los elementos y el recolector de eventos. Se encuentra implementada en la clase *Ttopologia*. Para llevar a cabo estas tareas, cada elemento debe utilizar un identificador que le será asignado por un generador de identificadores que posee la topología.

El aspecto primordial de la topología son los importantes métodos que implementa. Cabe destacar

tres de ellos, los cuales implementan los algoritmos de encaminamiento:

- *obtenerIPSalto()*, recibe como parámetros la dirección IP origen de un paquete y la dirección IP destino a donde se desea enviar y devuelve la IP del siguiente nodo de la ruta, por el que hay que seguir.

- *obtenerIPSaltoRABAN()*, este método realiza la misma labor que el anterior, pero calcula la dirección IP según el algoritmo *RABAN* (Routing Algorithm for Balanced Active Networks)[7].

- *obtenerIPSaltoRABAN()*, también realiza la misma labor, pero en este caso admite otro parámetro que es la dirección IP de un nodo adyacente al origen por el que no se desea pasar. A efectos prácticos, este método devuelve el segundo mejor salto posible para llegar del origen al destino.

4.2 El escenario de simulación

El *escenario* es una clase que contiene todo lo referente a un entorno completo de simulación. Reloj, recolector, topología, generadores de identificadores y de IP, nodos, enlaces, etc., se encuentran en un objeto de tipo *Tescenario*, de forma que es un solo objeto el que almacena todos los elementos de cada escenario. Algunos de sus atributos más importantes son:

- Una topología, de tipo *Ttopologia*, donde se almacenan todos los elementos y donde se realizan todas las conexiones entre ellos, como ya se ha comentado en el apartado 4.1.

- Un objeto de simulación, de tipo TSimulacion, que es donde se encuentra el recolector de eventos de simulación.

El método más importante del escenario es *generarSimulacion()*, que pone en funcionamiento el reloj de la topología y de este modo la simulación comienza a funcionar. En realidad, *Tescenario* incluye todos los ingredientes necesarios para funcionar aunque no exista interfaz de usuario.

5 Conclusiones

El presente trabajo propone el empleo de OpenSimMPLS como herramienta de innovación docente en asignaturas de redes y comunicaciones, justificado por el creciente interés que está despertando la tecnología MPLS. Asimismo, se ha demostrado que el simulador es una herramienta apta para la validación de resultados en proyectos de investigación relacionados con MPLS. Por tanto, con filosofía multiplataforma y licencia de software libre, se pretende potenciar su evolución, al recibir e incorporar todas las sugerencias que de su uso se deriven de una forma cooperativa.

Referencias

[1] Jose L. Marzo, Eusebi Calle, Caterina Scoglio, and Tricha Anjali, “QoS Online Routing and MPLS Multilevel Protection: A Survey,” IEEE Communications Magazine, October 2003.

[2] M. Kodialam and T. V. Lakshman, “Restorable Dynamic QoS Routing,” IEEE Communications Magazine, June 2002.

[3] E. Rosen et al., “Multiprotocol Label Switching Architecture,” RFC 3031, January 2001.

[4] Janus Gozdecki, Andrzej Jajszczyk, and Rafal Stankiewicz, “Quality of Service Terminology in IP Networks,” IEEE Communications Magazine, March 2003.

[5] G. Ahn, W. Chun, Design and Implementation of MPLS Network Simulator. Chungnam National University of Korea, February 2001.

[6] Análisis de la Integración entre Tráfico IP y Redes MPLS. Simulador MPLS. Miguel Ángel Martín Tardío, Miguel Gaspar Rodríguez, José Luis González-Sánchez. III Jornadas de Ingeniería Telemática. JITEL’01. Barcelona, 19-21 septiembre 2001.

[7] An Architecture to Provide Guarantee of Service (GoS) to MPLS. A. M. Domínguez-Dorado, F. J. Rodríguez-Pérez, J. L. González-Sánchez, J. L. Marzo, A. Gazo. IV Workshop in G/MPLS Networks. Girona. April 21-22, 2005.

[8] G. Ahn, W. Chun, Simulator for MPLS Path Restoration and Performance Evaluation, Chungnam National University, Korea, April 2001.

[9] MPLS Simulator: http://www-entel.upc.es/xavierh/mpls/

[10] OpenSimMPLS: http://patanegra.unex.es/opensimmpls/web/es/indiceES.html

---

Ú Este trabajo está financiado, en parte, por la Consejería de Educación, Ciencia y Tecnología de la Junta de Extremadura, Proyecto AGILA, con código No. 2PR03A090 y por la CICYT, Proyecto MAIN-RA, con código No. TIC2003-05567.

PAFET4: Un ejercicio de predicción sobre la innovación en servicios telemáticos

Juan C. Dueñas, Vicente Burillo, José L. Ruiz
Departamento de Ingeniería de Sistemas Telemáticos. Universidad Politécnica de Madrid
ETSI de Telecomunicación. Ciudad Universitaria sn.
28040– Madrid
Teléfono: 913366831 Fax: 913367333
E-mail: jcduenas@dit.upm.es, vbm@dit.upm.es, jlruiz@dit.upm.es

Abstract*. PAFET4 is a research initiative, fostered by AETIC, COIT and the Ministry of Industry, that tries to foresee the professional competences required by future Information Technologies staff. The domain of research of PAFET4 has been focused on the technical innovations that will reach the market in medium-term, restricted to the area of telematic services and management of digital contents. The prospecting of the services that may be successful in the future depends of technical areas of development, and also of the identification of the general business and market conditions that allow it. The work presented here is underway, which leads to preliminary conclusions; also the methodology of the research is somehow different to the traditional research practices in telematic engineering. Anyway, the authors think that an explanation of the work it is being performed may contribute to understand the business and market conditions driving telematic engineering research.*

1 Introducción

El proyecto PAFET4 es una iniciativa de investigación industrial promovida por la Asociación Española de empresas de Tecnologías de la Información y las Comunicaciones (AETIC), el Colegio Oficial de Ingenieros de Telecomunicación (COIT) y el Ministerio de Industria, y llevada a cabo por personal del Departamento de Ingeniería de Sistemas Telemáticos de la ETSIT de la Universidad Politécnica de Madrid.

El estudio, cuyos resultados preliminares se presentan, es el cuarto de una serie de proyectos con el nombre PAFET (Propuesta de Acciones para la Formación de Profesionales de Electrónica, Informática y Telecomunicaciones). El primero de ellos trataba de determinar las necesidades de profesionales TIC en un escenario económico y tecnológico óptimo (antes de la "burbuja .com"); el segundo estudio se dirigió a la búsqueda de nuevos nichos de empleo relacionados con las TIC, y el tercero identificó un sector "transformador de la tecnología", intermedio del sector tradicional TIC compuesto por operadores y fabricantes, y los usuarios finales (particulares o corporativos). En todos los casos han sido investigaciones que tienen que ver con la prospección tecnológica, las implicaciones de mercado y su repercusión en las necesidades de profesionales del ámbito de las TIC y, por lo tanto, en su formación.

La elección del dominio de trabajo de PAFET4 –los servicios telemáticos y el ámbito de los contenidos digitales- viene a reflejar una situación característica del entorno tecnológico de las TIC, que refuerza el papel de la ingeniería telemática en el marco general de las TIC. En este caso, el foco de interés consiste en estudiar qué innovaciones TIC, en cuanto a servicios y contenidos digitales, alcanzarán el mercado a medio plazo, cuáles de ellas sobrevivirán en ese entorno, y cual será su impacto en la formación de profesionales TIC (el estudio no se dirige de forma única a la formación universitaria).

El tipo de investigación a la que hay que hacer frente en un estudio como este es multidisciplinar: en primer lugar es necesario conocer los mecanismos que gobiernan la innovación tecnológica y cuáles son las tendencias o iniciativas más importantes en este sector. También es necesario conocer la lógica del mercado en el cual se mueven las empresas, organismos públicos, entidades y personas. Aunque este es un entorno de trabajo tradicionalmente transparente para el ámbito científico-técnico, el mundo corporativo –incluidos los departamentos técnicos- se ve influenciado cada vez más por los métodos y elementos del mercado. Por último, es preciso conocer el sistema de formación disponible en la actualidad, para poder ajustar el grado de cambio en cuanto a contenidos que deben sufrir los planes de estudios de los futuros profesionales TIC.

Lejos del ánimo de los autores el saberse capaces de conocer en detalle esos tres ámbitos de conocimiento de las TIC (el sistema de innovación y las innovaciones potenciales, el mercado, el sistema de formación), por lo que, en el desarrollo del estudio, se ha recurrido a técnicas de investigación comunes en otros ámbitos de la industria (y más cercanas a las ciencias sociales): identificación de preceptores, realización de entrevistas, recopilación bibliográfica, revisiones de cartografías tecnológicas, validación por expertos, etc. Para el lector técnico es preciso indicar que se está tratando de predecir el futuro,

labor en la que normalmente no se aplica plenamente el método científico-experimental y se encuadra dentro de las técnicas de toma de decisiones.

El presente documento ofrece un resumen preliminar de las conclusiones del estudio, puesto que éste se encuentra todavía en fase de realización. El documento presenta en la sección 2 una visión de los resultados generales obtenidos en los estudios previos PAFET, la metodología, presupuestos y objetivos perseguidos en PAFET4 en la sección 3, algunas conclusiones preliminares en cuanto a la estructura de los mercados TIC en la sección 4, y reseña algunas innovaciones tecnológicas de potencial en el futuro en la sección 5. El documento acaba con algunas conclusiones generales acerca de la política científica, industrial y académica aplicables al ámbito de los servicios y contenidos digitales.

2 El punto de partida: PAFET 1, 2, 3.

2.1 PAFET 1

El primero de los estudios “Propuesta de Acciones para la Formación de Profesionales de Electrónica, Informática y Telecomunicaciones” [1] se dirigía a las empresas del sector de las telecomunicaciones. Este primer estudio, como ya se ha comentado, se dedicó a analizar el déficit de profesionales TIC cualificados y la necesidad de evolución de conocimientos técnicos. El estudio se produjo antes del desplome del mercado (burbuja de las .com); ya en ese momento se identificaron lagunas evidentes en la formación de los profesionales TIC. Estos elementos vinieron a reforzar las ideas propuestas por otras iniciativas de ámbito europeo como Career-Space.

Entre los resultados de detalle más relevantes se puede mencionar la identificación de un área de conocimiento mixta entre la telecomunicación clásica y la informática. De hecho, la integración de departamentos de investigación y docencia universitarios del ámbito de la ingeniería eléctrica con la informática es una recomendación expresada por los expertos académicos de Career-Space. Seguramente este resultado no sorprenderá a los investigadores y docentes del área de la ingeniería telemática; lo que sí sorprendió a los autores del estudio es encontrar esa recomendación emitida en el ámbito europeo.

Otro elemento de singular interés que se observó en ese momento, es que las empresas del sector ya estaban iniciando un proceso de reducción de personal en las áreas de desarrollo, mantenimiento y producción; el crecimiento de empleo para los profesionales TIC se podría producir en las funciones de la empresa relacionadas con la comercialización. Este cambio en la orientación profesional exige titulados con competencias personales (aparte de las técnicas).

2.2 PAFET 2

El proyecto PAFET 2 se denominó “Evolución de los perfiles Profesionales TIC en la Sociedad del Conocimiento” [2], y trató de profundizar en las posibilidades de evolución de los perfiles profesionales TIC, en España y su dimensión internacional. En esos momentos se comenzaba la discusión sobre la modificación curricular en las Universidades (proceso de Bolonia – Espacio Europeo de enseñanza superior), y se exploraba la implicación empresarial (Career-Space) [3][4] en la enseñanza superior.

El estudio se limitaba a proponer y evaluar varios esquemas de implantación del Espacio Europeo de enseñanza superior (lógicamente aplicables a los estudios superiores de telecomunicación e informática) que siguen siendo objeto de análisis y discusión en los ámbitos de política académica. En cuanto a las conclusiones del estudio ya se hacía evidente la necesidad de renovar el modo de trabajo de las escuelas y facultades universitarias y adaptar los contenidos docentes para, entre otras mejoras, dar cabida a la formación en competencias personales y empresariales.

2.3 PAFET 3

En el estudio PAFET 3 “Perfiles Emergentes de Profesionales TIC en Sectores Usuarios” [5] se daba por conocida la situación de reducción del volumen de empleo de las empresas tradicionales dentro del sector de las TIC, y se trataba de descubrir nuevos yacimientos de empleo.

Estos yacimientos de empleo se comenzaban a vislumbrar como asociados al ámbito corporativo (con la identificación del profesional de infraestructuras TIC, y del planificador TIC asociado a la dirección de la empresa). También se identificó como un dominio de actividad económica creciente el dominio del ocio y doméstico, además de los dominios cercanos al ámbito de la sanidad, tanto en sus vertientes de sector público como privado.

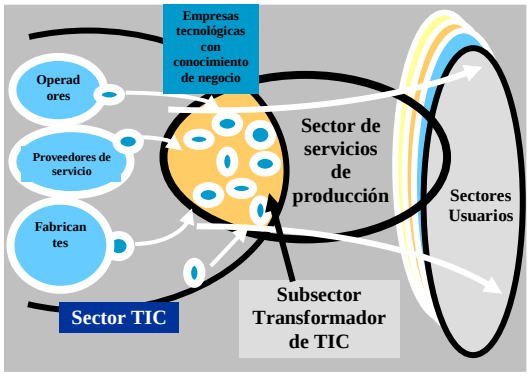


Fig. 1: el subsector transformador de las TIC.

En dicho estudio también se identificó la situación emergente de un nuevo sub-sector transformador de

la tecnología, formado por pequeñas empresas o departamentos de las empresas tradicionales en el sector, y cuyo objetivo fundamental era el de transformar la tecnología que las grandes empresas innovadoras eran capaces de poner en el mercado "en bruto", y adaptarlas a nichos de uso o dominio específicos. Los agentes del sector transformador establecen un puente entre los proveedores "duros" de la tecnología y los usuarios, trabajan con márgenes de tiempo pequeños, necesitan de innovación, y conocen bien el dominio de la aplicación de la tecnología.

Este sector transformador de la tecnología forma el embrión de lo que posteriormente en el estudio PAFET4 hemos identificado como "cadenas de valor" de los servicios.

También es pertinente indicar que cada uno de estos estudios, con objetivos diferentes, concluía con la identificación de los perfiles profesionales emergentes asociados a los hallazgos correspondientes. La metodología de cada trabajo ha ido adaptándose a las necesidades del estudio y a los elementos de conocimiento disponibles, pero hay que indicar que en todos ellos, los elementos metodológicos han tenido por objetivo extraer información oculta o difusa y por lo tanto no abordable desde el punto de vista científico tradicional (bien por medio de entrevistas, bien mediante encuestas, paneles, etc).

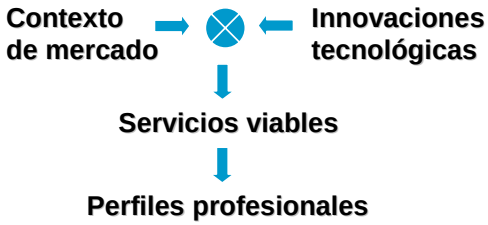


Fig. 2: esquema metodológico de PAFET4.

2.4 Objetivos de PAFET 4

El estudio PAFET4 "Soporte a la implantación de los servicios y contenidos digitales", que estamos llevando a cabo y cuyos resultados preliminares estamos presentando tiene los siguientes objetivos:

- Identificación de nuevos servicios telemáticos (incluyendo los relacionados con la gestión de los contenidos digitales) a medio plazo.
- Valoración con respecto a su capacidad de despliegue y utilización
- Identificación de las competencias necesarias en los profesionales TIC

Estos objetivos toman como punto de partida que los servicios telemáticos y los contenidos digitales seguirán creciendo en importancia económica, y que existen innovaciones tecnológicas relativas a ellos que pueden llegar al mercado.

3 Elementos metodológicos

3.1 Metodología del trabajo

Los elementos metodológicos básicos pueden observarse en la Fig. 2. Respecto a las actividades de soporte del estudio, se mencionan aquí únicamente para mostrar que son actividades de investigación industrial utilizadas, bien conocidas y aceptadas en otros ámbitos de la ingeniería. Estas son:

1. Estudio documental del dominio.
2. Diseño y realización de entrevistas a expertos del dominio.
3. Elaboración de los resultados preliminares.
4. Panel de expertos: evaluación y contraste de los resultados preliminares.
5. Elaboración de conclusiones.

En cualquier caso, es preciso disipar una sombra de duda sobre la posibilidad de otras aproximaciones al problema, utilizadas en otros estudios por este mismo equipo de trabajo: la aproximación más objetiva de recogida de datos (numéricos) de campo, análisis estadístico de los mismos, elaboración y propuesta de hipótesis y su validación resulta muy adecuada para el estudio de fenómenos existentes, pero imposible cuando se trata de hacer una predicción acerca de las innovaciones tecnológicas, al menos por dos motivos:

- Se está tratando de analizar hechos en el futuro y no es posible demostrar que la extrapolación de los valores en el pasado pueda dar lugar a resultados válidos para el futuro.
- No existe una fuente fiable y coordinada de datos con la granularidad suficiente para nuestro estudio.

3.2 La innovación

Uno de los puntos de partida del estudio es la consideración de la innovación como elemento fundamental para comprender las TIC en la actualidad. Aquí, la innovación tecnológica se entiende como los cambios con base en la tecnología, ya sea en un uso nuevo de la existente o en el uso de una nueva tecnología. La innovación tecnológica se impone como una realidad cada vez más amplia (que incluye elementos tradicionales de investigación, desarrollo, innovación técnica, innovación de procedimientos), y se implanta con velocidad creciente.

Esta innovación se produce no solamente en el sentido convencional del término como resultado de una investigación científica planificada; en realidad cada vez más se produce como una innovación de

mercado planificada con criterios mercantiles, e incluso innovaciones espontáneas fuera de los circuitos científicos o industriales convencionales.

3.2 El mercado

El mercado es la estructura básica que engloba a los diferentes agentes en la explotación, gestión, comercialización y uso de las TIC –en nuestro caso, de los servicios y contenidos digitales-. Es la existencia de estos mercados la que, en último extremo, justifica el desarrollo de una infraestructura de formación. También es la evolución de estos mercados, tanto en lo microeconómico como en las magnitudes macroeconómicas, la que va a permitir fomentar o reducir la existencia de innovaciones.

3.3 La confluencia de mercado e innovación

Nuestro estudio considera que, dadas unas innovaciones tecnológicas potenciales en el ámbito de los servicios y contenidos digitales, sólo cuando se puedan explotar con éxito, se traducirán en creación de empleo, y pueden requerir de nuevos perfiles profesionales o modificaciones a los existentes.

Esta confluencia de las innovaciones y de las condiciones del mercado se puede analizar en primer término mediante un modelo conceptual, ya existente en otros ámbitos de la industria, denominado “cadena de valor”.

4 Ecosistema de valor de los servicios TIC

En esta sección se describe el modelo conceptual de cadena de valor de los servicios TIC, para posteriormente observar cómo existen varias fuerzas y fenómenos que nos llevan a trascender el modelo básico de cadena de valor hasta convertirlo en un modelo más rico y complejo que se viene denominando de forma genérica “ecosistema de valor de los servicios TIC”·

4.1 Los servicios TIC

En primer lugar es preciso definir el concepto de servicio TIC. Existen varias definiciones técnicas y también del ámbito de la economía y los negocios. De entre las técnicas, podríamos definir servicio como “función de un sistema”, aunque en el ámbito en el que estamos haciendo el trabajo tal definición es incompleta sin la presencia de la red, así que la definición completa es: “función de un sistema a la cual se puede acceder de forma remota”, o dicho de otro modo, el lugar del uso es diferente del lugar de la realización. En términos aún más técnicos, se puede hablar de la “implementación del servicio” como el lugar o elemento físico en el cual se ejecuta el servicio, “interfaz del servicio” como los medios y mecanismos que permiten al usuario, utilizando el protocolo de uso del servicio, acceder al servicio, “usuario” como aquel que usa el servicio a través de la interfaz anterior, y evidentemente, “red” como elemento físico-lógico que permite la interconexión entre los elementos indicados.

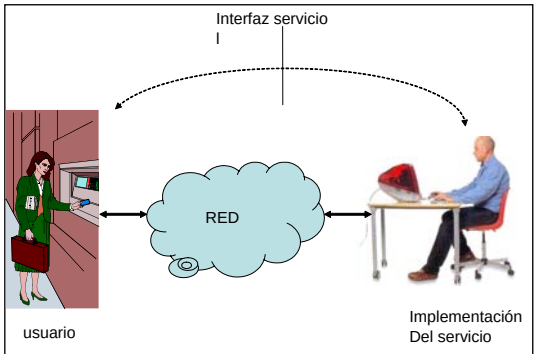


Fig.3: elementos de los servicios.

En cuanto a las definiciones desde el punto económico o de negocio, indicar que el término servicio aparece en las transacciones económicas como contrapuesto a “bien” y en una primera aproximación, necesariamente superficial, nos permitiría hablar de “elemento sobre el que se tiene la capacidad de uso o usufructo, pero no se tiene la propiedad”. También es de especial importancia indicar que cuando se habla de servicio se habla de “contrato por prestación de servicio”, y no de compra.

Así pues, y tratando de conjugar todos los elementos indicados hasta el momento, circunscribimos nuestro estudio a los servicios, definidos como aquellas funciones accedidas a través de una red, en las que los usuarios tienen el derecho de uso, pero la propiedad la mantiene el prestador del servicio (o servidor). La calidad de “remoto” de los servicios, o de accesible desde diferentes sitios y situaciones, implica la existencia de varias funciones (que se traducen a diferentes agentes), cada uno de los cuales cumple una o varias funciones en esta cadena de interrelaciones.

En la medida en la que el esquema conceptual descrito permite el acceso a los contenidos digitales (datos en diferentes formatos) al igual que a los servicios, nos encontramos en una situación en la cual tanto servicios como contenidos se pueden manejar de forma homogénea. En nuestro estudio estamos encontrando que las funciones generales del sistema para el acceso a los servicios son prácticamente las mismas que para el acceso a los contenidos digitales, y que las tecnologías y el entorno de mercado para unos y otros es cada vez más parecido (fenómeno para el que se suele utilizar el término “convergencia”). En el resto del documento nos referiremos a los servicios incluyendo el acceso a los contenidos digitales.

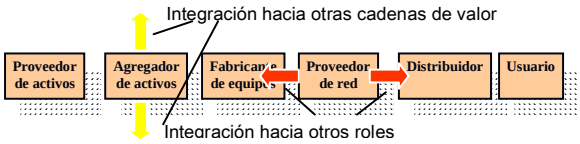


Fig. 4: cadena de valor de los servicios.

4.2 La cadena de valor de los servicios

De forma prácticamente directa de las definiciones anteriores llegamos a la conclusión de que son necesarias varias funciones para poder llevar a cabo un servicio. Un modelo conceptual tradicionalmente aceptado para reflejar estas situaciones es el denominado "cadena de valor". La cadena de valor es un modelo que viene del mundo industrial –particularmente de la fabricación-, y refleja el hecho de que diferentes agentes deben de interactuar, aportando valor, hasta alcanzar al usuario o consumidor. Una característica consustancial al concepto de cadena es la concatenación "lineal" de los agentes que van añadiendo valor hasta que se alcanza al usuario final.

En el caso de los servicios, y por la propia definición de éstos, encontramos inevitable la existencia de acuerdos o cadenas de agentes (empresas, por ejemplo), aportando valor o cumpliendo las diferentes funciones necesarias hasta alcanzar al usuario o consumidor. La Fig. 4 muestra un modelo abstracto de cadena de valor, compuesta por agentes, también abstractos, fruto de la síntesis resultante como producto del estudio y análisis de varias de estas cadenas de valor relacionadas con los servicios TIC:

1. Generador/ Proveedor de activos. Tanto para los servicios telemáticos como para los contenidos digitales, es preciso crear éstos. Con la palabra activos englobamos a ambos. El generador o proveedor se encarga de las fases previas a la explotación de los activos. Ejemplos de dichos proveedores pueden ser productores de software de servicios, agencias de noticias, casas discográficas, etc. En general, son todos aquellos agentes cuyo objetivo es el de crear y empaquetar activos digitales.

2. Agregador de activos. En general los activos digitales no son únicos; es más, su explotación comercial exige componer activos digitales diferentes (canales de televisión en canal de distribución).

3. Fabricantes de dispositivos y equipamiento. La provisión de los elementos físicos para la cadena de valor de los servicios es otra de las grandes funciones -que puede afectar de manera transversal a dicha cadena-. Piénsese, por ejemplo, en los servicios de telefonía móvil en los que existen fabricantes de dispositivos terminales con grandes nichos de mercado y visibilidad como marcas, pero también proveen del equipamiento de red necesario para desplegar y gestionar ésta.

4. Proveedor de servicios de comunicación. El proveedor de servicios de comunicación es el tradicional operador de red, sea esa fija, móvil, de datos… o de transporte. De forma histórica encontramos que estos proveedores tienen un activo esencial desplegado o por desplegar, pero que exige una inmovilización de activos económicos y con plazos de amortización largos.

5. Distribuidor-canal de comercialización. Es un rol difícil de identificar separadamente para los servicios y contenidos, pero en esencia, se trata de aquel que hace la oferta comercial y por lo tanto tiene capacidad de cobro (cuando se habla de la factura única o de la cuádruple oferta de telecomunicación se está identificando este rol de forma precisa).

6. Usuario final. ¿ Debería de aparecer este agente en la cadena de valor? ¿ es posible identificar un único agente con este nombre?. No tenemos una decisión concluyente a este respecto, pero en cualquier caso, tanto en el ámbito del negocio como en el tecnológico, comienza a ser evidente el papel activo que juegan los usuarios en la penetración de los servicios, por contraposición a modos de análisis tradicionales que consideran al usuario final único fuera del sistema de análisis, como condición de contorno.

En nuestro análisis hemos encontrado:

1. Que sólo se pueden ofrecer servicios telemáticos cuando aparecen una gran parte de los agentes mencionados,

2. Que para poder ofrecer servicios de forma coherente y eficaz, además de estable, es necesario que exista una condición mínima de comunicación o relación entre agentes que se denomina "interoperabilidad",

3. Que los servicios sólo van a tener éxito cuando se produce un equilibrio entre los diferentes agentes para la provisión de esos servicios, siendo el papel del usuario final en este equilibrio, determinante e influenciado por factores derivados del contexto social, económico y cultural.

4. Que el equilibrio será estable cuando los modelos de negocio internos de cada uno de los agentes sean estables.

4.3 El ecosistema de valor

También hemos encontrado, en nuestras entrevistas, paneles y análisis, que el concepto de cadena de valor se encuentra en revisión, está desbordado, porque no es capaz de representar las complejas relaciones que se dan entre los agentes. Entre las mayores limitaciones que presenta se pueden citar: que es un modelo estático, lineal y que no ofrece movimientos o caminos de realimentación, por lo que cualquier desequilibrio "local" en uno o más agentes tiende a propagarse por la cadena, por lo que el equilibrio de esta, con frecuencia, no puede ser estable.

Reconocemos el interés del modelo cadena de valor, para proponer un modelo que lo trasciende y mejora, aunque a costa de introducir una complejidad de modelado que, acercándolo a la realidad, puede hacerlo inmanejable.

Analicemos por un momento los tipos de inestabilidades o movimientos que pueden aparecer en una cadena de valor:

- Movimientos horizontales: un agente en una cadena de valor puede tratar de extenderse en la propia cadena de valor, ocupando varias funciones o roles. Se trataría de aquellas cadenas de valor en las que alguno de los roles no tiene competencia o cuyas funciones vayan quedando obsoletas (fenómeno de "comoditización") y exige al agente la asunción de roles cercanos para poder mantenerse en el mercado. Un efecto de esta integración horizontal se da también en cadenas de valor en los que hay agentes muy dominantes (o únicos), como por ejemplo en las cadenas de valor de los servicios de telefonía.

- Movimientos verticales. Este tipo de integración ocurre cuando un agente que cumple una función en una cadena de valor puede cumplir la misma u otra similar (tendríamos por tanto un movimiento "diagonal") en otra cadena de valor cercana. Este es el fenómeno que entendemos ocurre con la "convergencia" tecnológica, en la que, por ejemplo, un fabricante de terminales de servicios móviles también puede fabricar terminales para televisión digital; o un proveedor de servicios móviles puede proveer servicios en cadenas de banda ancha. Es necesario reconocer que, hoy en día, este tipo de integración vertical, facilitado por la difusión de la tecnología y los estándares abiertos es una gran fuerza que efectivamente puede hacer inservible el concepto de cadena de valor para hacer emerger el de ecosistema de valor.

El modelo de "ecosistema de valor" que estamos comenzando a explorar podría describirse como un modelo del mercado de los servicios telemáticos, en el cual se superponen varios tipos de cadenas de valor (para dominios diferentes de servicios), varias cadenas de valor para cada servicio, varios agentes para cada función o rol, relaciones de n a m entre agentes y roles, y evoluciones en estas relaciones. Se trata de un escenario ciertamente complejo en el cual identificamos tres grandes fuerzas motoras (la diferenciación en calidad o costes como estrategia de negocio, los aspectos sociales –incluyendo regulación, legislación, presiones de usuarios-, y la innovación tecnológica), representadas en la Fig. 5.

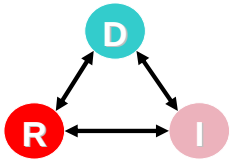


Fig. 5: fuerzas diferenciación-regulación-innovación.

De cara a nuestro estudio, por ahora concluimos que para que una innovación tecnológica llegue al mercado y necesite de nuevos profesionales TIC es preciso que se produzca una evolución del ecosistema de valor (creación de una nueva cadena de valor, modificación de alguna existente, convergencia entre cadenas de valor colindantes, etc.). Las condiciones mínimas para la "mutación" en el ecosistema tienen que ver con el reajuste del sistema diferenciación-regulación-innovación tras la aparición de una innovación tecnológica.

5 Innovaciones tecnológicas

Al tratar de la innovación tecnológica, hay cierta tendencia a pensar en nuevas tecnologías, en la creación y desarrollo de estas, produciendo cierta prevención en el mundo no tecnológico, cuya opinión generalizada es la de aprovechar las tecnologías disponibles -que no son pocas- enfriando la generación y desarrollo de otras nuevas, sobre todo por su elevado coste (investigación básica) y la insuficiente amortización de las existentes.

Nosotros, tecnólogos en el terreno de las TIC, cuando nos referimos a la innovación tecnológica en el contexto de los nuevos servicios, estamos pensando en los nuevos usos de las tecnologías disponibles con el fin de proveer esos nuevos servicios o mejorar los presentes. Evidentemente, la innovación tecnológica, vista desde esta perspectiva, surgirá de los nuevos usos de las tecnologías maduras e irá a remolque de las nuevas tecnologías que se vayan desarrollando.

No se trata, por tanto, de crear nuevas tecnologías, ni tan siquiera de desarrollar más las existentes. El asunto está en utilizar las tecnologías maduras y, si es adecuado, desarrollarlas en la medida conveniente. En cualquier caso, estos desarrollos tecnológicos no son objetivo de la innovación tecnológica que aquí interesa, aunque no se deben perder de vista las

tecnologías emergentes, que llegadas a un mínimo nivel de madurez, podrán sostener nuevas innovaciones tecnológicas orientadas a la provisión de servicios.

Aunque puede parecer sutil, esta consideración determina la aproximación que estamos llevando a cabo en nuestro estudio, con respecto a las fuentes de la innovación tecnológica: en lugar de proponer innovaciones, resulta más interesante (y productivo), analizar los diferentes mapas (cartografías) de innovación existentes en nuestro entorno. Concretamente hemos tomado como punto de partida la cartografía tecnológica realizada por el grupo ITEA [6], y en cuyo proceso de elaboración hemos participado los miembros del grupo de trabajo.

5.1 Innovaciones tecnológicas habilitadoras

Las futuras innovaciones tecnológicas, a las que nos venimos refiriendo en este trabajo, previsibles a día de hoy pueden contextualizarse en las tendencias tecnológicas que se vislumbran como marco de desarrollos futuros viables, de las que forman parte tanto los servicios como los contenidos.

Parece haber cierta coincidencia en considerar áreas de servicios y contenidos con viabilidad futura las siguientes:

- servicios desarrollados en la esfera Internet
- servicios provistos por las comunicaciones móviles
- servicios asociados a las plataformas de televisión digital interactiva
- servicios en el ámbito público y corporativo
- contenidos digitales, su formato y las plataformas de distribución

En estas áreas se pueden identificar las innovaciones tecnológicas en servicios y contenidos que harán viables a estos últimos. Los nuevos servicios pueden ser ordenados y agrupados en dominios, adoptando una perspectiva de uso de los mismos. Los dominios considerados en este trabajo han sido el doméstico-personal, el público-corporativo y nómada-móvil, coincidiendo con otros estudios de previsión tecnológica.

En el dominio doméstico-personal, dentro del área de Internet, se prevén los servicios de voz sobre IP, los p2p, los de acceso a la información (noticias, anuncios comerciales, anuncios administrativos, …), ocio, formación, e-comercio y otros más especializados. Los servicios móviles en este dominio se centran en los servicios básicos telefónico, SMS y servicios de localización. En el ámbito de la televisión digital interactiva, los servicios de ocio, información y publicitarios destacan sobre otras posibilidades.

El dominio público-corporativo incluye los servicios de banda ancha dirigidos a las empresas, los servicios web, los e-servicios (x2y), los servicios de infraestructuras (provisión de “hosting”, “housing”, redes privadas virtuales, etc), consultoría y transformación-adaptación de tecnología, e-administración y e-gobierno, formación y e-learning, información corporativa y administrativa, servicios de seguridad, acreditación y autentificación, entre otros.

El nomadismo y la movilidad confieren características propias a los servicios, tanto desde el punto de vista del usuario (ubicuidad del servicio) como desde la perspectiva de su provisión. La ubicuidad e instantaneidad de la provisión es un activo diferencial, que tiene sus limitaciones en la calidad y el ancho de banda alcanzables. Ejemplos de servicios provistos en este dominio, además de los servicios básicos, son los de información (noticias, información turística y de transporte, financiera, ..), servicios de localización, juegos, ocio (música, descarga de tonos de llamada e iconos, ..) entretenimiento de adultos, directorios móviles, mensajería, emergencias, …

Asociadas a estos servicios aparecen problemáticas específicas, cuyas soluciones son el terreno abonado para las innovaciones habilitadoras. Estas no son ajenas a los modelos de negocio, el mercado y, en general, el contexto que hace un servicio viable. Uno de los elementos clave en este es la existencia de las capacidades profesionales que hagan posible la concepción y el desarrollo de soluciones, generalmente innovadoras, que darán ventaja en el mercado a los productos o servicios que las incorporen.

Tabla 1: importancia relativa del rol en el dominio

Rol-dominio	**Doméstico**	**Público**	**Nómada**
Generador proveedor	**	**	*
Agregador	**		*
Fabricante equipos	**		**
Proveedor de comunicación	*	**	**
Distribuidor comercializador	**	*	
Usuario	**	**	**

6 Conclusiones preliminares

La tabla 1 presenta muy someramente los resultados a los que estamos llegando acerca de la importancia de cada uno de los roles identificados en los dominios de servicios seleccionados (el dominio doméstico, el dominio corporativo-público, el dominio nómada). La información incluida en la tabla refleja de forma aproximada el equilibrio de fuerzas previsto en el futuro para cada uno de los dominios. Con una

información tan abstracta –y ciertamente discutible en sus aspectos metodológicos, puesto que se ha obtenido a partir de entrevistas con los expertos identificados- es posible realizar algunas predicciones: en la introducción de innovaciones tecnológicas de servicios en el ámbito doméstico, es preciso coordinar varios roles, donde el proveedor de comunicación podría tener un peso relativo menor que otros roles, lo cual también permite indicar que existe una fuerza para que los agentes en este rol asuman otros roles (la integración horizontal de la que hemos hablado anteriormente). También en el ámbito de los servicios domésticos, los aspectos de comercialización y gustos del consumidor (usuario) van a ser determinantes.

En el dominio corporativo-público, el generador-proveedor de servicios y contenidos, junto con el proveedor de red, son los agentes con mayor capacidad de influencia –dejando aparte el usuario final-. En el dominio nómada es en el que van a tener mayor influencia los roles de fabricantes y operadores de red.

La tabla 2 presenta también resultados provisionales acerca de las innovaciones tecnológicas que según los expertos TIC y la revisión documental –incluyendo las cartografías tecnológicas- tendrán mayor difusión e impacto en cada uno de los dominios de servicios.

Tabla 2: importancia relativa de la innovación en el dominio.

Innovación-dominio	Doméstico	Público	Nómada
Gestión y protección de contenidos	**		*
Activos FOSS	*	**	
Usabilidad y accesibilidad	**		**
Interoperabilidad, middleware, servicios web	*	**	
Personalización y adaptación al contexto	**	*	**
Seguridad y gestión de la identidad	*	**	*

Las innovaciones (se trata de áreas de innovación) identificadas como más prometedoras son:

- Gestión y protección de contenidos: todas aquellas innovaciones que van orientadas a facilitar la creación, adaptación, manejo, empaquetamiento, aseguramiento de derechos (DRM *Digital Rights Management*) y explotación de contenidos digitales, especialmente multimedia, incluyendo nuevos formatos.

- Activos FOSS (*Free Open Source Software*): los métodos, herramientas, plataformas y modelos de cooperación del mundo del código abierto son áreas de innovación que se prevé crezcan en el futuro y que afecten a los agentes del mercado.

- Accesibilidad y usabilidad: innovaciones centradas en la percepción de la calidad de los servicios y contenidos digitales por parte de los usuarios finales. Ante un escenario futuro de "sociedad de la información", el éxito del despliegue de los nuevos servicios pasa por aumentar su potencial de uso, no solamente para usuarios entrenados, sino también por personas de edad avanzada y discapacitadas.

- Interoperabilidad, "middleware" y servicios web: se trata –al igual que las anteriores- de un área de innovación muy activa, y se centra en los aspectos de los servicios que tienen que ver en último término con la visión arquitectónica expresada por el W3C.

- Personalización y adaptación al contexto: estas innovaciones tienen como referente las propuestas de "inteligencia ambiental" y "sistemas ubicuos" en los que los servicios están fuertemente determinados por su contexto de uso (contexto físico, geográfico, cultural, legal, de usuarios).

- Seguridad y gestión de la identidad: como un área de innovación emergente se han identificado todas las técnicas que tienen que ver con la construcción del "mundo virtual" que incluye información sobre identidades personales y corporativas.

El trabajo al que nos hemos referido en esta exposición se encuentra en su fase final; esperamos, en los meses hasta su conclusión, avanzar en la elaboración y validación de estos resultados preliminares, y formular las conclusiones en cuanto al impacto que las innovaciones en servicios telemáticos y de gestión de contenidos digitales tendrán en el empleo y por tanto en la formación de los futuros ingenieros. Entretanto, comienzan a aparecer publicaciones del ámbito tecnológico con un enfoque similar al aquí expresado [7].

Referencias

[1] G. León, M. Gamella, C. Matías, F. Sáez, J. C. Dueñas, A. Bernardos. PAFET1 - Propuesta de acciones para la formación de profesionales de electrónica, informática y telecomunicaciones para las empresas del sector. ANIEL, COIT, 2001. ISBN: 84-609-0014-2.

[2] G. León, A. Bernardos, V. Burillo, J. C. Dueñas, C. Matías, F. Sáez. PAFET2 - Evolución de los perfiles profesionales TIC en la sociedad del conocimiento. ANIEL, COIT, 2002. ISBN: 84 - 609-0015-0.

[3] Generic ICT Skills Profiles. 2001. Consorcio Career-Space.

[4] ICT Curricula for the 21st century- Curricula guidelines. Consorcio Career-Space. 2001.

[5] V. Burillo, J. C. Dueñas, A. Bernardos, C. Matías. PAFET3 - Perfiles emergentes de profesionales TIC en sectores usuarios. ANIEL, COIT, 2004. ISBN: 84 - 609-0016-9.

[6] ITEA Technology Roadmap for Software-Intensive Systems. 2004. Oficina ITEA-Information Technology for European Advancement.

[7] R. Saracco, Leveraging Technology and the Market to Bring Telecommunications Business into the Next Decade. Global Communications Newsletter, IEEE Communications, Febrero 2005, 43/2.

Laboratorio de Interconexión de Redes Telemáticas

N. Rodríguez[1,2], N. Cañamares[1,2], P. Bustamante[1,2], E.Reina[1], R. Zubillaga[2], M. A. Orea[2]
[1] CEIT y [2]TECNUN (Universidad de Navarra)
Manuel de Lardizábal 15, 20018 San Sebastián
Teléfono: 943 21 28 00 Fax: 943 21 30 76
E-mail: nrodriguez@ceit.es

Abstract*. Telecommunications market is in constant evolution towards new networks that support new applications and services. In this context an original and innovative lab where students have to configure network equipment from physical to application level is presented. This lab allows Telecommunication Engineering students to be close to telecommunication market and to adopt the different roles that companies can play (internet service providers, contents providers, different kinds of clients…) Students replicate a real network situation to analyze different kind of networks in depth in order to solve the problems related to each type of network technology. In addition, this paper includes a way of managing a lab composed by several technologies configured in a limited set of network equipments to obtain the best performance test bed.*

1 Introducción: Motivación y objetivos del laboratorio

El sector de las telecomunicaciones se enfrenta a enormes retos derivados del proceso de convergencia entre los sectores de telecomunicaciones, audiovisual e Internet. Las necesidades actuales de los usuarios requieren nuevas redes, servicios y aplicaciones capaces de integrar información multimedia. En este contexto el panorama del mercado de las telecomunicaciones aparece como un entorno de redes diversas, interconectadas entre si, sobre las que se ofrecen variados servicios. Es un mercado complejo que combina agentes proveedores de red, de servicios y de contenidos para acercar unas aplicaciones de calidad al usuario final con independencia de la red y del medio físico empleado.

La demanda de servicios de comunicación actual es creciente e implica una mejora continua de las prestaciones de los sistemas de telecomunicaciones, que se han convertido en un reto científico-tecnológico permanente. En este marco, el laboratorio de interconexión de redes ha sido diseñado como un entorno versátil y configurable para reproducir las variadas combinaciones de tecnologías de red y aplicaciones que conviven en el mercado actual. Este banco de pruebas de estructura modificable resulta idóneo para el desarrollo, testeo y monitorización de aplicaciones en las que la red juega un papel decisivo. Entre estas aplicaciones destacan aquellas que garanticen parámetros de calidad de servicio a usuarios conectados a través de diferentes redes o aplicaciones distribuidas entre computadores en paralelo, físicamente distantes, que suman sus capacidades de procesamiento para resolver algoritmos complejos.

El laboratorio, aunque es una herramienta de investigación cualificada que resulta muy útil para la realización de proyectos fin de carrera, tesis doctorales y cursos de postgrado, constituye además una herramienta docente de gran potencial. Entre las diferentes asignaturas que pueden cursarse en el laboratorio, en este artículo se presenta Modelado y Dimensionado de Redes Telemáticas como demostrador de la versatilidad y aplicabilidad de este laboratorio de investigación al campo docente.

2. Modelado y Dimensionado de Redes Telemáticas

La asignatura Modelado y Dimensionado de Redes Telemáticas es una asignatura eminentemente práctica impartida en el último curso de la titulación de Ingeniería de Telecomunicación, cuyo objetivo principal es mostrar la implementación real de los conocimientos teóricos adquiridos en las aulas, a lo largo de las asignaturas de telemática de la titulación de Ingeniería de Telecomunicación (como “Redes de Telecomunicación” y “Redes, Sistemas y Servicios de Comunicaciones”). En el laboratorio los alumnos diseñan, instalan, configuran y gestionan todas las redes presentes en el laboratorio, desde el nivel físico y aprendiendo a solucionar problemas de forma estructurada: conectividad, nivel de enlace, nivel de red, problemas derivados de la seguridad y el acceso, y nivel de aplicación.

Se fomenta y valora la autonomía de los alumnos, impulsando la familiarización con las herramientas de ayuda que proporcionan los fabricantes para análisis y detección de fallos en la conexión y a optimizar las búsquedas de información en función del problema concreto a resolver.

Como aspecto colateral, pero no menos importante, esta asignatura proporciona a los alumnos la posibilidad de familiarizarse con la problemática que llevan consigo los diferentes tipos de redes adoptando

los puntos de vista de los distintos agentes implicados en esta área proveedores de red, proveedores de servicio, proveedores de contenido y clientes.

3. Metodología y estructura

3.1 Estructura física del laboratorio

El laboratorio en el que se imparte la asignatura está formado por un parque heterogéneo de equipos de interconexión de redes que permiten reproducir la situación actual del mercado de las comunicaciones en el que conviven tecnologías de generaciones diferentes. En este conjunto se encuentran desde equipos con capacidad de procesamiento pequeña, pero suficientes para dar conectividad a redes de área local de pymes, hasta equipos ubicados en el backbone de una red. Esta variedad resulta imprescindible para la comprensión global del proceso de evolución e integración de redes y tecnologías.

El material del laboratorio está compuesto por 10 PCs con diferentes prestaciones y una estación de trabajo (incluyendo tanto servidores como clientes) y 10 routers de distintos modelos, con versiones de IOS y tarjetas de red variadas. A modo de ejemplos extremos, el router más pequeño es modelo 801 de CISCO [1] con un puerto Ethernet y otro RDSI y el que tiene mayor capacidad de interconexión de redes es un 3640 de CISCO [2] con 4 puertos RDSI, 2 puertos serie, 4 ATM, 2 FXS y 2 analógicos, que constituye el backbone de interconexión de redes WAN y LAN. Para completar el laboratorio se dispone de equipos ATM (un router 3640 y un Light Stream de CISCO [3]) y ADSL (un DSLAM [4] y un router) que acercan estas tecnologías hasta las redes locales. Además se dispone de una central telefónica que da señalización al conjunto y un Call Manager de CISCO con funcionalidades de central telefónica IP.

En la figura 1 se presenta una fotografía en perspectiva del laboratorio y en la figura 2 un plano en planta del laboratorio con el equipamiento mencionado. En esta figura se pueden diferenciar dos tipos de equipos: los que permanecen fijos en el armario y los móviles (casi todos los routers). Los routers se ubican en las mismas mesas de los PCs para favorecer el contacto más cercano del alumno con el equipo y facilitar la diferenciación de cables y puertos. Las mesas han sido cableadas hasta un panel de parcheado situado en el armario con los equipos fijos y la central telefónica para posibilitar el acceso desde los puestos a los equipos del armario.

En la figura 3 se incluye una fotografía de los equipos fijos del laboratorio incluidos en los armarios.

Figura 1: Laboratorio de interconexión de redes

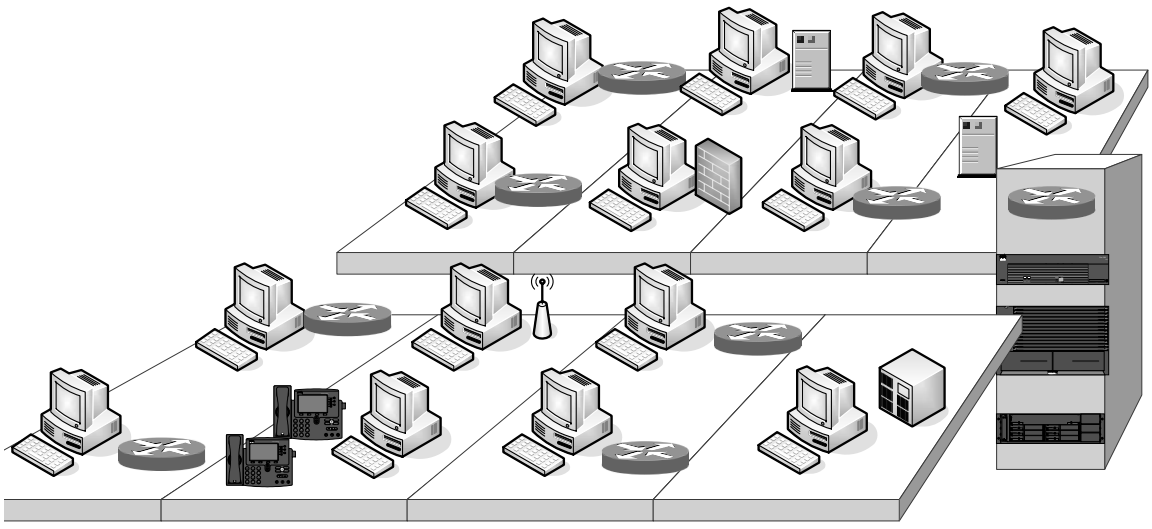
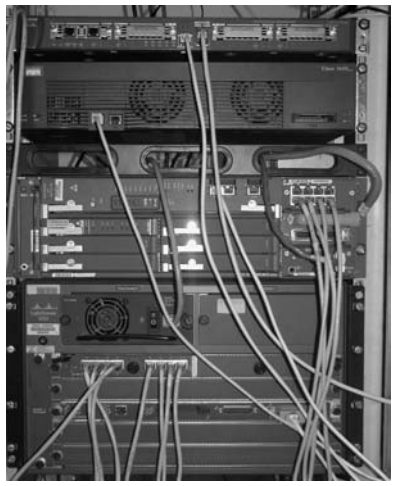
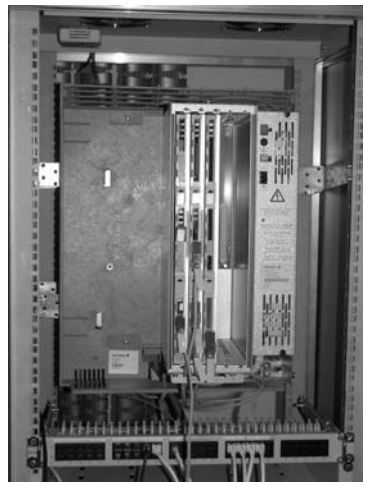


Figura 2: Plano en planta del laboratorio con equipos móviles y fijos



a) Backbone ATM y Centralita IP



b) Centralita telefónica

Figura 3: Detalles de los armarios de equipos fijos del laboratorio

3.2 Contenido teórico de la asignatura

La asignatura Modelado y Dimensionado de Redes Telemáticas, es fundamentalmente práctica y se imparte en un entorno de laboratorio. Sin embargo, es necesario un conocimiento teórico previo para maximizar el aprovechamiento de los alumnos. En la primera sesión introductoria, junto a conceptos básicos de direccionamiento, switching y routing, se presenta el panorama actual de las telecomunicaciones, las tecnologías implicadas en este campo y las tendencias actuales del mercado. Se hace un especial hincapié en la amalgama de tecnologías que compone la red actual ya que esta convivencia de protocolos y redes, equipos de última generación mezclados con tecnologías pendientes de sustitución ya que es el escenario habitual con el que se encuentran al incorporarse al mercado laboral.

Como complemento a la sesión introductoria y recordatorio de los conceptos adquiridos en otras asignaturas, se presenta una selección de enlaces de CISCO donde se incluye la bibliografía básica con información de las tecnologías involucradas en el laboratorio, los manuales de configuración de estas tecnologías, ejemplos de configuraciones y herramientas de detección de problemas (debugs). Estas páginas constituyen el contenido teórico de la asignatura. Cuando los alumnos aprenden a gestionar esta información disponen de una herramienta de gran valor para su futuro profesional en este campo que les hace capaces de configurar y solucionar la mayor parte de los problemas relacionados con la instalación de las redes telemáticas.

El listado de links se puede consultar en [5]

3.3 Prácticas desarrolladas

En el laboratorio se realizan dos tipos de prácticas diferentes: prácticas guiadas en las que se familiarizan con los conceptos básicos de redes y aprenden a configurar cada una de las tecnologías involucradas y una práctica final, que implica tanto el diseño como la instalación de una red, donde se combinan todos los conocimientos adquiridos durante la asignatura. En este punto se presentan las prácticas guiadas, mientras que la práctica final se considera con suficiente entidad para dedicarle el apartado 4.

Sin tener en cuenta las prácticas introductorias, las prácticas guiadas que se realizan en el laboratorio se pueden dividir en tres bloques diferentes: tecnologías de acceso, tecnologías de transporte y servicios complementarios.

- **Tecnologías de acceso**: El primer grupo de prácticas incluye aquellas en las que se aprende a configurar las redes de acceso: ADSL [6], RDSI [7], accesos telefónicos a través del modem y puntos de acceso wireless. En estos casos se configuran ambos extremos, el del proveedor y el del abonado.
- En el grupo denominado **tecnologías de transporte**, se incluyen aquellas prácticas en las que se configuran redes que conectan redes entre sí: ATM [8], Frame Relay [9], X-25 [10] y un radio enlace por microondas.
- **Servicios complementarios:** En este tercer grupo se incluyen prácticas que complementan a las anteriores como configuración de una central telefónica que permita tanto abonados analógicos como

conexiones de datos a routers RDSI, configuración de redes privadas virtuales (VPN) [11] y configuración de un firewall. En este grupo, se incluyen también algunos servicios como la instalación y configuración de un servidor de video y la configuración de una central IP [12].

En la figura 4 se incluye la práctica de RDSI

3.4.- Estructura lógica de las prácticas

Se dispone de conjunto de equipos de red con prestaciones completamente diferentes en cuanto a procesador, tipo y número de puertos, sistema operativo y software de aplicación. Otro factor a tener en cuenta es que se tienen 8 grupos de prácticas simultáneos y al no ser equipos de características idénticas, los alumnos no pueden realizar la misma práctica a la vez. Para gestionar la rotación de prácticas se realiza un circuito individualizado para cada grupo. La rotación de prácticas supone tener en cuenta varios factores:

- Hay prácticas "llave" para otras. Por ejemplo para realizar "Frame Relay con backup por RDSI" se presupone "Configuración de RDSI".
- Dos prácticas que configuren el mismo equipo no son compatibles. "Frame Relay" y "Enlace X-25"
- Prácticas que necesitan señalización (RDSI) son incompatibles con la configuración de la central telefónica.
- Prácticas sobre el mismo canal físico son incompatibles: "Configuración de Radioenlace por microondas" y "Configuración de LAN Wireless"

Este punto, que a priori puede parecer trivial, constituye la clave que posibilita la realización de las prácticas con un número de equipos muy diversos. En la rotación juegan un papel importante las prácticas denominadas complementarias ya que en casi todas ellas se utilizan equipos específicos que no interfieren con la configuración de las redes (firewall, servidores, centralita IP).

4.- Práctica final

La última práctica de la asignatura, que supone la mitad de los créditos asignados, es una práctica no guiada. En ella se propone una situación que real en la que intervienen diferentes agentes: proveedores de servicio, proveedores de contenido y clientes distribuidos en diferentes lugares e interconectados por diferentes tecnologías. A través del guión se presenta el panorama actual del mercado de las telecomunicaciones en el que cada uno de los usuarios finales involucrados tiene diferentes necesidades.

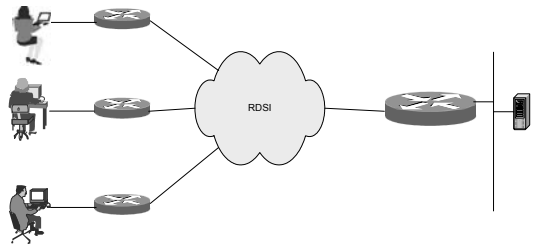


Figura 4: Esquema de la práctica de RDSI

A modo de ejemplo ilustrativo se presentan las características generales de la práctica final propuesta durante el curso 04-05, así como la topología de red diseñada e implementada por los estudiantes con los equipos disponibles en el laboratorio.

- **Empresa de contenido**, llamada MII-Europa, que ofrece video bajo demanda y difusión de eventos en vivo. Esta empresa tiene sedes en Bruselas y Luxemburgo, con dos redes locales en las que comparten datos. En el momento en que se plantea el problema, esta empresa tiene una persona desplazada en una feria en París y un directivo de viaje que se puede conectar a la red interna de la empresa desde el hotel vía VPN. Ambos tienen conexión de voz y datos.
- **Empresa proveedora de servicio**, TECNUN-Telecom, a quien pertenecen las conexiones punto a punto, y las redes de telefonía analógica, RDSI y ATM-ADSL que interconectan las diferentes ciudades en las que opera MII Europa y las ciudades de los clientes
- **Empresa cliente**, R&R, que contrata cursos de formación on line y que tiene dos sedes en edificios de la misma ciudad muy próximos, conectados con un radio enlace de microondas. En las oficinas de R&R hay mucho movimiento de puestos y no existe cableado fijo.
- **Cliente particular**, que quiere conexión puntual para eventos bajo demanda con un modem analógico.

Entre los requisitos del problema se incluyen planes de contingencia y de backup, servicios de telefonía IP para reducir costes entre sedes y garantías de seguridad en la empresa proveedora de contenido, separando las redes públicas con contenidos audiovisuales de las redes privadas de la empresa. Los alumnos, repartidos en equipos que forman las diferentes empresas, deben reproducir la situación planteada y se les evalúa en función de los objetivos conseguidos tanto en situación de funcionamiento normal de la red y en situación de caída de las líneas principales.

Antes de acometer la configuración de la red en el laboratorio, cada uno de los grupos propone la solución que estima más adecuada a la situación propuesta, con las restricciones impuestas por el material del que se dispone en el laboratorio. Se selecciona aquel diseño de topología de red que, a priori, tiene mayores expectativas de viabilidad y se encarga la dirección del proyecto al grupo que la ha desarrollado. Este grupo se encarga de coordinar la numeración RDSI y los DLCI, distribuir las direcciones IP para las diferentes redes locales, asignar los números de circuitos virtuales ATM y coordinar las claves y contraseñas para los protocolos de autenticación. Otra tarea del grupo director es la coordinación del enrutamiento. Como primer resultado de este trabajo los alumnos fueron conscientes de la importancia de una dirección eficaz, asociada a la buena coordinación entre extremos para garantizar la comunicación. En la figura 5 se incluye el diagrama de la red configurada por los alumnos durante el primer cuatrimestre del curso académico 04-05 como ejemplo para mostrar la complejidad del trabajo abordado.

Al terminar de configurar la red se comprueba la conectividad entre todos los puntos y la calidad de los servicios ofertados por el proveedor de contenido: video, mail y web. Estos servicios se comprueban en situación de funcionamiento correcto de la red y en situación de backup por RDSI o ADSL cuando fallan los enlaces primarios. Como último punto de la práctica final, los alumnos proponen mejoras que contribuyan a incrementar la calidad y el número de los servicios ofertados por el proveedor. Entre las mejoras propuestas por los alumnos se incluyen sustituciones de equipos por otros de tecnología más avanzada, nuevas tarjetas en alguno de los equipos existentes, alternativas para incompatibilidades detectadas, variación de parámetros de configuración que mejoran la calidad de servicio…

5- Futuro de la asignatura

Con la experiencia adquirida tanto en el montaje del laboratorio como durante el periodo de docencia impartida, se proponen nuevas prácticas que se irán incorporando progresivamente en la docencia de este laboratorio.

La primera práctica que se incorpora en la próxima edición de la asignatura es la captura de paquetes en las diferentes redes incluidas en el laboratorio. El objetivo de esta propuesta es acercar a los alumnos los diferentes protocolos involucrados enseñándoles a diferenciar la secuencia de encapsulados que sufren los paquetes al cambiar de red y los campos

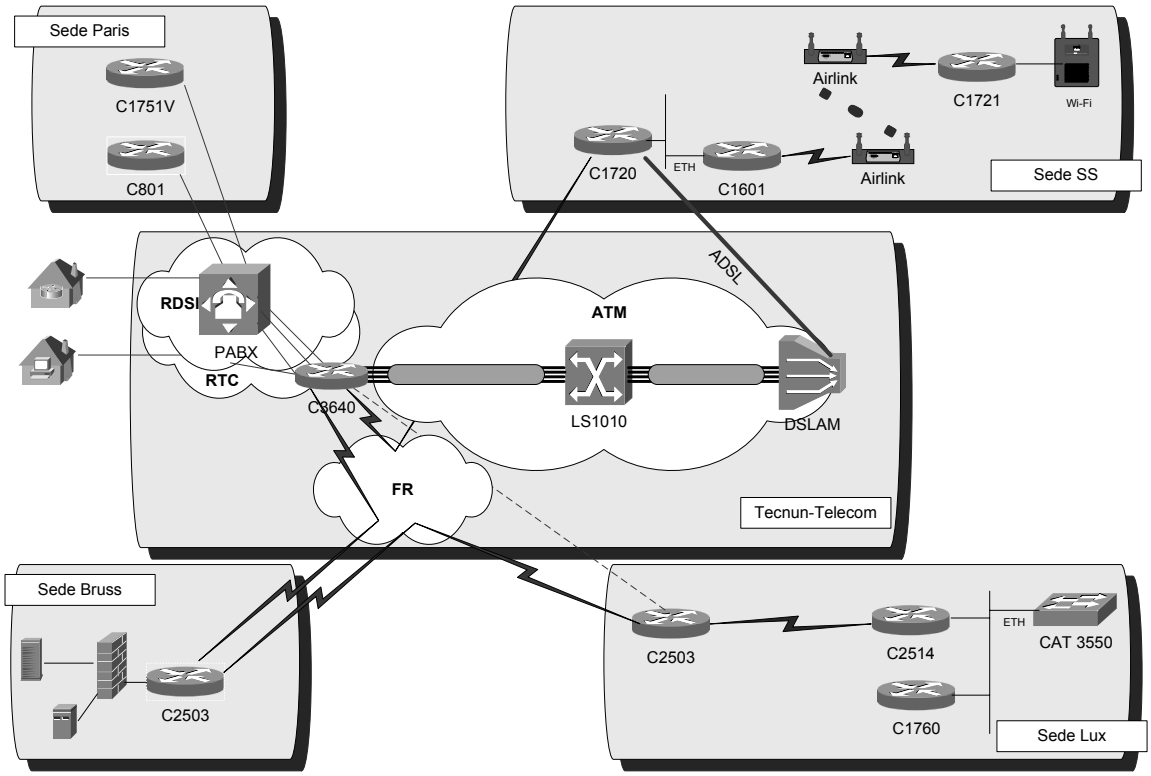


Figura 5: Diagrama de red diseñada por los alumnos del primer cuatrimestre del curso 2004-2005

incluidos en cada una de las tramas de una forma didáctica. Esto permite profundizar en conceptos como niveles, clases de servicio, priorización de tráfico en función de la aplicación y mostrar de forma práctica los problemas de mantenimiento de calidad de servicio a lo largo de la red que se derivan de la convergencia de redes y servicios en el entorno actual.

Otra de las prácticas a incorporar es la configuración y gestión centralizada de la red. La implantación de un sistema de monitorización permite que los alumnos conozcan la utilidad y el manejo de estas herramientas.

6.- Conclusiones

Al concluir el periodo de docencia de una asignatura de estas características y evaluar los resultados alcanzados, se pueden extraer conclusiones de diferente índole:

En cuanto a los *conocimientos de los alumnos* que han cursado la asignatura, se comprueba que los alumnos han afianzado las nociones básicas de redes (direccionamiento IP, routing, switching…) y los conocimientos de las diferentes tecnologías LAN y WAN que se adquirieron de forma teórica en otras asignaturas. Sin embargo, la importancia de la docencia de una asignatura de este tipo, no radica tanto en los conocimientos concretos como en las habilidades adquiridas: manejo de documentación, conocimiento práctico de las situaciones en que se emplea cada tecnología en función de los requerimientos y necesidades de conexión, y sobre todo, la importancia de la coordinación entre los agentes involucrados en el establecimiento de la comunicación.

En cuanto al *laboratorio* en sí, se concluye que es posible la implementación de un laboratorio diseñado con equipamiento muy diverso que combina tecnologías de interconexión y en el que no existen dos routers con las mismas características. Esta particularidad complica un poco la docencia ya que en el aula conviven simultáneamente grupos configurando tecnologías radicalmente distintas. Sin embargo, con una adecuada distribución de las prácticas, estableciendo un itinerario individual para cada uno de los grupos, es posible combinar los circuitos de prácticas para que todos los alumnos tengan la posibilidad de configurar todas las tecnologías.

Referencias

[1]http://www.cisco.com/en/US/products/hw/routers/ps380/tsd_products_support_series_home.html

[2]http://www.cisco.com/en/US/products/hw/routers/ps274/tsd_products_support_series_home.html

[3]http://www.cisco.com/en/US/products/hw/switches/ps1893/tsd_products_support_series_home.html

[4]http://www.cisco.com/en/US/products/hw/switches/ps298/tsd_products_support_series_home.html

[5]Web de la asignatura "Modelado y Dimensionado de Redes Telemáticas" http://www.tecnun.es/asignaturas/redtelema/pagina_5.html

[6]http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/dsl.htm

[7]http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/isdn.htm

[8]http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/atm.htm

[9]http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/atm.htm

[10]http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/x25.htm

[11]http://www.cisco.com/en/US/products/sw/iosswrel/ps1831/products_configuration_guide_chapter09186a00800d981f.html

[12]http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122limit/122z/122zj15/itsv30/

Experiencias en la utilización de *middleware* de código abierto para el aprovisionamiento de servicios extremo a extremo en el sector residencial

Jose L. Ruiz, Juan C. Dueñas, Manuel Santillán
Departamento de Ingeniería Telemática. Universidad Politécnica de Madrid
ETSI de Telecomunicación. Avda. Complutense S/N.
28040 – Madrid
Teléfono: 913366831
E-mail: jlruiz@dit.upm.es, jcduenas@dit.upm.es, santillan@dit.upm.es

Abstract*. Open Source Software (OSS) is changing the established paradigm and business models in the IT sector. While, in some environments, the usage of OSS is feasible –such as web servers, where Apache servers running on top of Linux machines are widespread-, its usage is still under debate for more complex scenarios, such as end to end (e2e) service provisioning for networks of residential gateways. A particular example of such challenges appears in the operation and management of thousands of machines, and the complexity derived from the heterogeneity of home services, networks and devices. In this article we will explain the solution that we are proposing to these challenges, based on the massive usage of OSS, middleware and standard communication protocols. The work presented here is still in progress and is being carried out within the context of the EUREKA ITEA Osmose project.*

1 Introducción

El papel que el código abierto (OSS-*Open Source Software*) puede jugar en la difusión y despliegue de los servicios telemáticos es actualmente objeto de debate en la comunidad de investigación y también en la de la práctica de la ingeniería. Los enfoques basados en el uso de componentes de código abierto tienen sus raíces en el mundo de los estándares abiertos que tanto han contribuido a conformar la situación actual de la ingeniería telemática.

Tras unos años en los que la situación del sector económico relacionado con los servicios telemáticos ha atravesado malos momentos, parece que estamos asistiendo a una nueva etapa de crecimiento y despliegue. En esta situación de creación de nuevos mercados y despliegue masivo de los servicios, el uso de soluciones tecnológicas de código abierto puede funcionar como catalizador de la innovación, impidiendo la formación de monopolios *de facto*.

Son muy diversas las razones que llevan a las empresas a tomar en consideración el OSS para sus modelos de negocio, entre ellas los elevados costes de las licencias de *software* propietario o la necesidad de independencia con respecto al proveedor del *software*. Siguen existiendo, sin embargo, algunos motivos que actúan como barreras a la hora de integrar este tipo de activos. Por mencionar algunos podríamos hablar de las garantías de mantenimiento que puedan obtenerse del proveedor, que en este caso es una comunidad de código libre, o la compatibilidad con los modelos de negocio ya establecidos en la empresa.

En el presente documento pretendemos explicar la viabilidad del uso de OSS en un caso de uso muy concreto, el aprovisionamiento de servicios extremo a extremo en redes de pasarelas residenciales. Este caso de uso constituye un escenario ideal para poner en práctica, en forma de demostrador, los resultados obtenidos en el proyecto de investigación europeo EUREKA ITEA Osmose, proyecto que tiene como principal objetivo el desarrollo de tecnologías *middleware* de fuente abierta, como motor de una ventaja competitiva europea para el desarrollo de sistemas abiertos. El contenido de este artículo repasa desde un punto de vista descriptivo los aspectos principales de la arquitectura del escenario, aunque en la mayor parte de los casos no trataremos al detalle cada uno de los elementos debido a las limitaciones del espacio disponible.

El orden de exposición que iremos siguiendo será el siguiente, en primer lugar presentaremos el escenario de despliegue sobre el que estamos trabajando, intentando identificar los principales actores involucrados y poner de manifiesto cuáles son los retos a los que nos enfrentamos. A continuación describiremos cómo la solución propuesta en el proyecto integra activos *middleware OSS* para la configuración de la arquitectura de referencia que da soporte al conjunto de funcionalidades deseadas. Antes de presentar las conclusiones del trabajo, dedicaremos una sección a la explicación de algunos de los servicios desarrollados para validar este estudio. Terminaremos con las conclusiones derivadas de este trabajo realizado como experiencia de validación de uso de OSS en un escenario complejo.

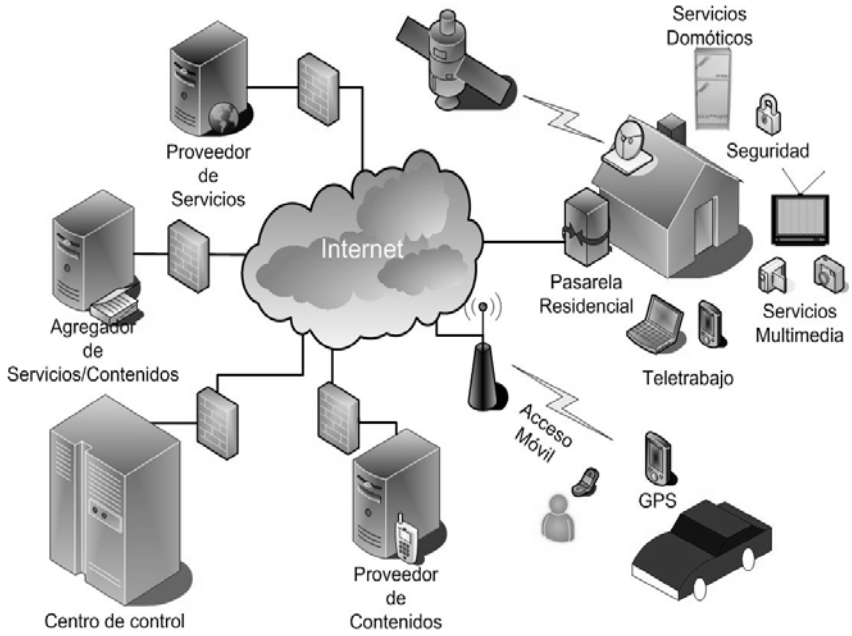


Fig. 1 Escenario de aprovisionamiento e2e

2 Escenario de aprovisionamiento

En esta sección presentamos el escenario sobre el cual se desarrolla el trabajo. Con el objeto de aclarar las ideas principales, hemos creado el diagrama conceptual que aparece en la Fig. 1.

Los actores esenciales involucrados en el escenario pueden identificarse en el diagrama. En uno de los extremos aparecen los proveedores de servicios y contenidos, cuyos modelos de negocio girarán en torno al aprovisionamiento de activos al sector residencial. Las posibilidades en este ámbito son múltiples: servicios de seguridad, servicios domóticos, vídeo bajo demanda, VoIP, grabaciones personalizadas de TV, etc.

En el otro extremo tenemos los hogares de los usuarios, en esta ubicación es destacable la presencia de la **pasarela residencial**. El papel de la pasarela residencial es de vital importancia en este escenario, porque constituye el elemento integrador entre los dispositivos y redes del hogar y el resto de los actores del escenario.

Es necesario distinguir en la pasarela residencial tres planos distintos desde el punto de vista lógico, como puede verse en la Fig. 2:

1. Plano de comunicaciones, la pasarela residencial dispone de un acceso de banda ancha a Internet, típicamente mediante xDSL o cable, aunque opciones inalámbricas como WiMax o GPRS/UMTS pueden ser necesarias en aquellos casos donde el despliegue de cable no sea viable por motivos económicos o técnicos. Este acceso de banda ancha se usa para canalizar tráfico generado o consumido por los dispositivos del hogar. La pasarela residencial dispone de interfaces de red que permiten interactuar entre sí y acceder desde un único punto desde fuera, a los dispositivos domésticos. Esta funcionalidad tiene una importancia especial debido a la gran diversidad de tecnologías de red que existen en el ámbito doméstico: IEEE 802.11x, Ethernet, Bluetooth, EIBus, Longworks, X10…

2. Plano de ejecución, en nuestro escenario una pasarela residencial no es simplemente un encaminador. La existencia de un plano de ejecución convierte al dispositivo en una plataforma de ejecución de servicios, o pasarela de servicios. Como veremos en el apartado siguiente la implementación del plano de ejecución está basada en tecnologías estándar. Esto permite a los proveedores minimizar el esfuerzo necesario para el desarrollo, despliegue y mantenimiento de sus servicios.

3. Plano de operación y gestión, este plano adquiere una especial relevancia en una pasarela residencial. Este plano tiene que ser capaz de gestionar tanto el plano de comunicaciones como el de ejecución. La gestión del plano de ejecución consiste en uno de los aspectos novedosos de nuestro trabajo, como veremos en la siguiente sección.

Uno de los objetivos que se persiguen con la introducción de la pasarela doméstica en el hogar es acercar las tecnologías de información y comunicaciones al conjunto de la sociedad, sin importar el nivel de formación técnica que puedan tener los usuarios. Una pasarela residencial debe ser

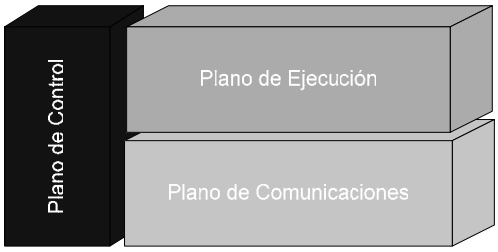


Fig. 2 Planos lógicos de una pasarela residencial

un dispositivo cuya gestión resulta completamente transparente para el usuario. Por la novedad del trabajo, en la siguiente sección describiremos de forma más extensa los planos de ejecución y de control de la pasarela.

El actor que aparece marcado como centro de control en la Fig. 1 será el encargado de ejecutar en lugar del usuario las operaciones de gestión y operación de las pasarelas.

3 Arquitectura del escenario

En esta sección vamos a cubrir las arquitecturas de dos de los actores principales que aparecen en el escenario: la pasarela residencial y el centro de control.

Desglosaremos la arquitectura de estas dos entidades analizando el tipo de soporte del que disponemos en términos de tecnologías *middleware*, plataformas y servicios. Iremos explicando las tecnologías según aparezcan , y las razones que justifican su aplicación.

3.1 Arquitectura de la pasarela

En primer lugar, describimos la solución desde el punto de vista de la infraestructura. De nuevo nos valdremos de un diagrama conceptual, ver la Fig. 3, para explicar las ideas principales. Intentaremos en la medida de lo posible hacer referencia a los planos lógicos de la pasarela doméstica (Fig. 2), y describiremos la figura desde abajo hacia arriba.

En primer lugar debemos destacar que vamos a utilizar como sistema operativo una distribución de Debian personalizada *ad hoc* [17], a la que se han añadido un conjunto de funcionalidades identificadas como requisitos básicos que un dispositivo que ha de funcionar como una pasarela doméstica deberá cumplimentar.

La elección de este sistema operativo de fuente abierta se justifica desde el punto de vista económico, ya que el despliegue de otras alternativas comerciales es mucho más costoso, pero también desde un punto de vista técnico. Las capacidades de personalización en cuanto a los componentes que conforman la distribución base del sistema operativo permiten

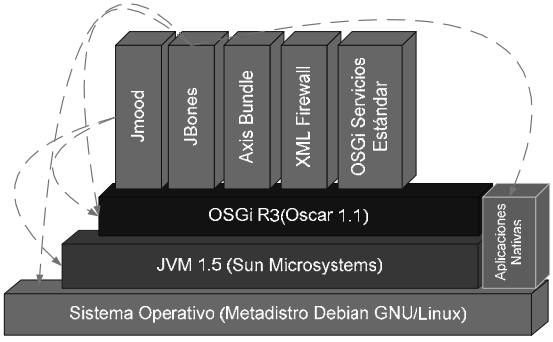


Fig. 3 Arquitectura de la pasarela

incluir en la distribución inicial todo el *middleware* necesario para la operación de la pasarela.

Otra de las ventajas de este sistema operativo son las herramientas para la gestión de paquetes que proporciona, que permiten la automatización de la infraestructura base de la distribución desde repositorios de *software* gestionados y/o autorizados por el centro de control. Esta capa forma evidentemente parte del plano de control de la pasarela doméstica.

Si continuamos hacia arriba en la figura nos adentramos en el núcleo del plano de ejecución que está compuesto por una plataforma OSGi [7], en particular la implementación de fuente abierta Oscar[18], ejecutando sobre una máquina virtual Java.

Dada la importancia de esta pieza de *middleware* en el escenario, merece la pena detenerse en explicar las características que nos llevan a utilizarla. La iniciativa OSGi fue lanzada en 1998 con el apoyo de más de 50 empresas del sector de los sistemas empotrados, operación de red, fabricantes de equipamiento, etc. Su objetivo consiste en desarrollar especificaciones para definir una plataforma de servicios Java, con capacidades para actuar como pasarela entre Internet y las redes de ámbito local, con lo cual se alinea perfectamente con los objetivos del escenario que estamos desarrollando.

Las especificaciones OSGi se basan en tecnologías Java, entre otras cosas por las ventajas de portabilidad en ejecución sobre distintas arquitecturas físicas de la pasarela. A grandes rasgos, la especificación está dividida en dos secciones: el entorno de ejecución (marcado como OSGi R3 en la Fig. 3) y la plataforma de servicios, definida como el conjunto del entorno de ejecución más una serie de especificaciones de servicios estándar. Este conjunto extra de servicios enriquece a las capacidades básicas de la plataforma con funcionalidades como la configuración de servicios, la gestión de permisos o la administración de usuarios.

La misión del entorno de ejecución consiste en asegurar la correcta ejecución y la interacción entre los servicios, mediante la implementación de las funcionalidades de registro, donde los servicios pueden ser dinámicamente dados de alta/baja o localizados por otros servicios.

La unidad de despliegue de servicios en OSGi (definidos éstos como interfaces del lenguaje Java) se conoce como *bundle*. Un *bundle*, es un archivo JAR (Java Archive) cuyo contenido puede estar compuesto por clases Java, bibliotecas nativas y cualquier otro tipo de ficheros (imágenes, archivos de texto...). Las aplicaciones en OSGi pueden estar compuestas por uno o más *bundles*, los cuales interactúan entre sí por medio de los servicios que van registrando en el entorno de ejecución.

La razón que nos han llevado a utilizar la plataforma OSGi para implementar el plano de ejecución de la pasarela residencial es el conjunto de capacidades de gestión remota que ofrece para la gestión dinámica del ciclo de vida de las aplicaciones [20], y también en cuanto al resto de los aspectos relevantes para el plano de gestión de la pasarela como puedan ser la gestión de permisos y la gestión de usuarios.

OSGi es, por tanto, el punto de partida y la base de la plataforma de ejecución de la pasarela, pero existen algunos aspectos que no quedan resueltos por OSGi desde el punto de vista de la especificación o que no hemos podido encontrar en las implementaciones de fuente abierta y que hemos tenido que implementar en el contexto del proyecto. El resto de componentes que aparecen ubicados por encima de OSGi están pensados para cubrir esos huecos:

JBones

La posibilidad de que las aplicaciones y servicios sean instalados dinámicamente sobre una plataforma de ejecución OSGi está considerada, pero nada se dice en la especificación sobre quién y cómo se deben controlar los procesos de despliegue de las aplicaciones, incluyendo todas las actividades de instalación y actualización.

Evidentemente esta tarea no puede dejarse en manos del usuario, por lo que la solución adoptada debía de ser necesariamente autogestionada o, alternativamente, remotamente gestionada.

Esta necesidad nos llevó a diseñar e implementar *JBones* [1][6][12], una pieza de middleware, que responde a las necesidades de despliegue en este contexto. Como aspectos novedosos de esta herramienta podemos destacar la gestión de despliegue de aplicaciones automatizada con resolución de dependencias entre componentes y el puente a la gestión de paquetes nativos, en el caso de que alguna aplicación a nivel OSGi necesite de algún componente de sistema operativo o nativo (por ejemplo, codecs de vídeo).

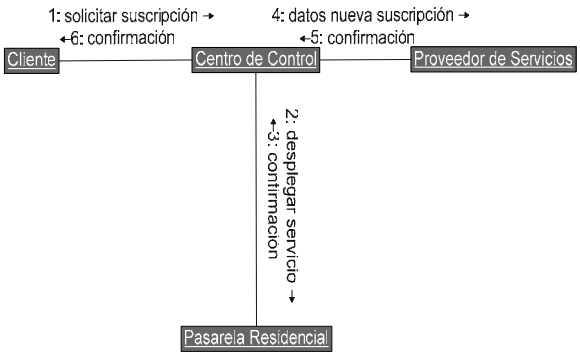


Fig. 4 Ejemplo de interacción *Web Services* entre los actores

JMood

La especificación de OSGi dedica un capítulo de su sección no normativa a indicar la necesidad de incluir un agente de gestión que se encargue de centralizar las actividades de gestión de la plataforma OSGi. Sin embargo, ninguna de las implementaciones abiertas de OSGi proporciona un agente de gestión.

JMood [13], como implementación del agente de gestión es otro de los componentes *middleware* que hemos añadido para implementar la arquitectura de la pasarela. *JMood* se basa en la tecnología JMX Java Management Extensions [3], ampliamente utilizada en otros contextos como el de gestión de aplicaciones de servidor J2EE.

JMood constituye la pieza fundamental del plano de gestión de la pasarela dado que, en integración con *JBones*, permite gestionar remotamente el plano de ejecución en su conjunto, incluyendo la plataforma de ejecución OSGi, los servicios estándares y todos los servicios proporcionados por terceros, mediante las interfaces de extensión que proporciona.

Axis Bundle

Se ha optado por utilizar los protocolos estándares relacionados con las tecnologías *Web Services* en el plano de comunicaciones. Todas las interacciones establecidas entre los actores del escenario (los proveedores de servicios, el centro de control y los instalados en la pasarela) se basan en comunicaciones SOAP/XML, podemos ver un ejemplo de esas interacciones en la Fig. 4.

El componente Axis *bundle* es el *middleware* que implementa las funcionalidades del plano de comunicaciones, tal y como representa la Fig. 2. Este *bundle* se basa en el motor de *Web Services* Apache Axis. El esfuerzo realizado en este elemento se ha centrado en dos aspectos. En primer lugar, empaquetar Axis para que pudiera ejecutarse sobre una plataforma OSGi y en segundo lugar, se han definido una serie de servicios y utilidades que permiten a los servicios de la pasarela OSGi publicarse como *Web Service* sin necesidad de tener que implementar los aspectos de comunicaciones por ellos mismos, sino basándose en las infraestructuras proporcionadas por el *bundle*.

XML Firewall

Las necesidades de seguridad en el acceso a la pasarela doméstica, tanto desde el centro de control para realizar actividades de gestión, como por el propio usuario de la pasarela, para accesos externos al hogar han obligado a implementar mecanismos de seguridad del nivel de aplicación mediante el filtrado de los mensajes XML/SOAP que se intercambiarán con las pasarela durante las habituales tareas de autenticación y autorización. El XML *firewall* [5] es el componente middleware que se encarga de este aspecto.

Servicios OSGi estándar

Debido a que algunos de los servicios no tenían implementación en las comunidades de fuente abierta ha sido necesario crear implementaciones para los servicios de configuración, *TIDuma* [15] y de gestión de permisos, *TIDpema* [16].

3.2 Arquitectura del centro de control

Una vez que ya hemos explicado la arquitectura de la pasarela vamos a pasar a tratar la arquitectura del centro de control. No entraremos tanto en detalle como en la sección anterior, dado que este aspecto se abordó con mayor profundidad en [4].

Aunque en un dominio no habitual, la arquitectura del centro de control responde al patrón de lo que se ha dado en conocer como sistemas de soporte a la operación. El punto diferenciador de este nodo se da en la focalización en las actividades necesarias para dar a conocer un servicio, desplegarlo en el entorno de ejecución, configurarlo y dejarlo en un estado inicial estable.

La arquitectura del centro de control, como puede observarse en la Fig. 5 aglutina un conjunto de funcionalidades complejas, algunas de las cuales están disponibles en componentes comerciales, como es el caso de los centros de atención al cliente. Estas funcionalidades se pueden desagregar en cinco bloques: despliegue de servicios, gestión/atención de usuarios, gestión de pasarelas (monitorización, configuración y mantenimiento), gestión de los proveedores de activos y mantenimiento del repositorio de aprovisionamiento.

Hemos tomado como referencia los resultados de OSS/J [19] como marco conceptual en la definición de los bloques funcionales del centro de control. El resultado es un centro de control diseñado mediante una arquitectura multicapa basada en componentes

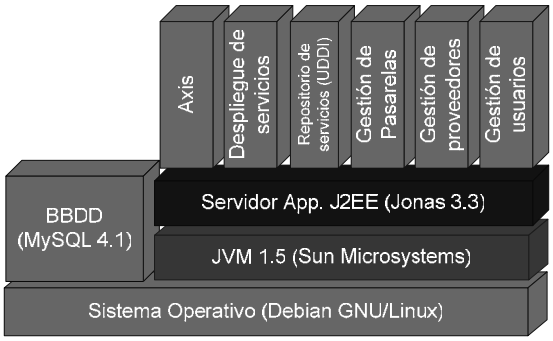


Fig. 5 Arquitectura del centro de control

reutilizables y tecnología de contenedores aprovechando la infraestructura proporcionada por Jonas, un servidor de aplicaciones J2EE de código abierto, sobre el cual se despliegan cada uno de los servicios descritos anteriormente. Esta infraestructura convierte al centro de control en un sistema de alta disponibilidad y capaz de soportar grandes cargas de trabajo –situación similar a la de los servidores de gestión de red convencionales.

Merece la pena hacer referencia a otros dos elementos que aparecen en la figura y que no se han comentado anteriormente: el motor de servicios Web y el UDDI. El motor de servicios web está encargado de publicar y permitir interacción entre los servicios desplegados sobre el centro de gestión, los proveedores de servicios y las pasarelas residenciales, todo ello utilizando las tecnologías de servicios web comentadas anteriormente. El UDDI será el registro para todos estos servicios. La capa de persistencia del centro de control se proporciona por una base de datos relacional.

4 Validación del escenario

Una vez que hemos cubierto todos los aspectos de infraestructura, pasemos a su validación. Adicionalmente a las pruebas habituales: unitarias y de integración, se decidió que la mejor manera de validar el escenario a nivel de sistema era mediante el desarrollo de servicios prototipo que permitieran validar el escenario de extremo a extremo.

Con este objetivo en mente se han desarrollado en el contexto del proyecto servicios que abarcan diversos ejes de funcionalidad: TV-IP interactiva o de control domótico. Dado que no es posible describir con la suficiente amplitud cada uno de los servicios desarrollados hemos escogido BarkIDS [**14**], servicio de seguridad para el hogar, como referencia para ilustrar los aspectos principales.

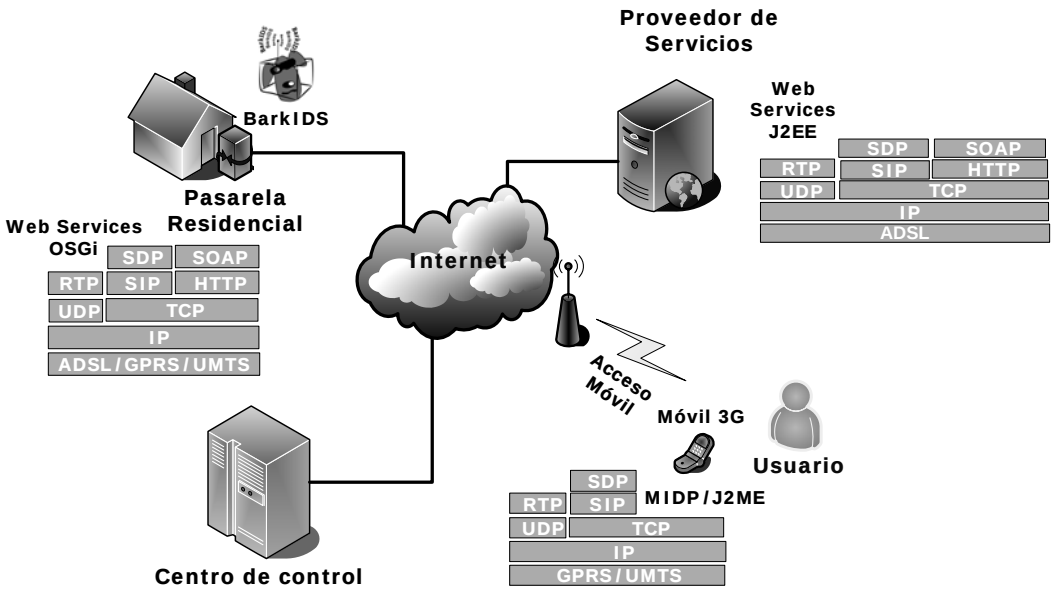


Fig. 6 Tecnologías y protocolos BarkIDS

Los servicios de vigilancia y seguridad para el hogar son actualmente instalados, controlados y administrados por empresas especializadas que emplean dispositivos habitualmente propietarios y creados *ad hoc*: cámaras de vigilancia, centros de control de vídeo, centros de grabación, etc. Sin embargo, BarkIDS no requiere de hardware especializado ni de agentes externos para gestionarlo, ya que gracias a su naturaleza esencialmente *software* puede ser gestionado remotamente de manera automática o por el propio usuario, y puede emplear cualquier tipo de cámara de vídeo digital para monitorizar el hogar y/o enviar alarmas al usuario ante la detección de un movimiento.

Este servicio es capaz de conectarse a una o más videocámaras para, mediante técnicas de detección de movimiento, alertar al usuario, esté donde esté, cuando se produzca una intrusión en su domicilio. Adicionalmente el usuario tiene la posibilidad de acceder al flujo de vídeo capturado por las cámaras desde cualquier lugar para observar el estado de su vivienda. Del mismo modo el servicio se puede utilizar para la vigilancia de mascotas, monitorización de menores o ancianos, etc.

Para intentar explicar la arquitectura general del servicio BarkIDS hemos incluido en la Figura 6 las entidades principales: el usuario, la pasarela doméstica, el proveedor del servicio y el centro de control. Adicionalmente en la figura también pueden observarse cuales son las tecnologías y torres de protocolos utilizados.

De esta manera podemos observar que el servicio está formado básicamente por dos componentes, uno que se despliega sobre la pasarela del usuario y que se encarga de detectar y enviar las alarmas detectadas por movimiento y otro que se ubica en la sede del proveedor del servicio, que se encargará de centralizar la recepción de esas alarmas y reenviarlas apropiadamente al usuario. El proceso de despliegue del servicio en la pasarela del usuario se realiza valiéndose de la infraestructura *middleware* de la pasarela y de las funcionalidades del centro de control. Todo usuario que disponga de una pasarela doméstica conectada a una cámara podrá darse de alta en el centro de control y una vez cursada la suscripción (ver Figura 4), para lo cual se utilizan las infraestructuras *middleware* ya mencionadas en el apartado anterior, se producirá de forma automática el despliegue del servicio a la pasarela del usuario.

5. Mapa tecnológico

Una vez descrito el escenario del conjunto de tecnologías involucradas hemos creído conveniente adjuntar una tabla en la que se resumen los elementos principales de los componentes utilizados en el escenario.

Dado que el trabajo realizado en el proyecto ha seguido los modelos de colaboración característicos de las comunidades de fuente abierta, indicaremos los autores involucrados en desarrollo de los componentes, destacando en negrita al autor principal en cada uno de los casos. Seguiremos el siguiente código numérico como leyenda de la tabla: Telefónica I+D (1) [8], Telvent (2) [9], DIT-UPM(3), ObjectWeb(4) [10] y la Apache *Software* Foundation(5)[11], aunque esta última no es parte del proyecto EUREKA ITEA Osmose.

Nombre	Descripción	Licencia	Autores
Jonas	Servidor J2EE	LGPL	**4**
Axis	Motor Web Services	ASL	**5**
Oscar	Plataforma OSGi	LGPL	**4**
Axis Bundle	Motor *Web Services* para pasarelas domésticas	GPL	**2**, 3
JMood	Agente de gestión de pasarelas JMX	LGPL	**3**
JBones	Motor de despliegue de *software* para pasarelas	LGPL	1, 2, **3**
Juddi	Servidor UDDI	ASL	**5**
TIDuma	Implementación de OSGi user admin.	GPL	**1**
TIDpema	Implementación de OSGi permission admin	GPL	**1**
XML Firewall	Seguridad XML	GPL	**2**, 3
JVM 1.5	Máquina virtual Java	**Propietaria**	**Sun Microsystems**
BarkIDS	Servicio de seguridad por videovigilancia	LGPL	**3**

Todos los componentes *software* mencionados en la tabla se han publicado con un modelo de licencia que permite su libre distribución, e incluso en el caso de que se deseara podrían crearse productos derivados de ellas y distribuirlos con otros modelos de licencia (no en el caso de los productos con licencia GPL-*General Public License*). Sin embargo, existe un riesgo identificado en cuanto a la libertad de la distribución de máquinas virtuales Java. Las máquinas virtuales de libre distribución no están habitualmente en un estado de desarrollo tan avanzado (en cuanto a versión de la plataforma Java soportada) y maduro como sus alternativas comerciales, con lo que deberían asumirse los costes de estas licencias o renunciar a distribuirlas junto con la pasarela.

6 Conclusiones

En el artículo hemos presentado, a modo de resumen, el trabajo que hemos realizado en el proyecto EUREKA ITEA Osmose [21] como esfuerzo básicamente de integración, pero también de investigación y desarrollo de componentes middleware para el aprovisionamiento de servicios en un contexto residencial.

Los resultados de la validación realizada hasta el momento son bastante positivos. Hemos sido capaces de reutilizar y crear el escenario al completo mediante *software* abierto, lo cual posibilitaría por un lado modelos de negocio en los que los costes por licencias no fueran una barrera insalvable, permitiendo la entrada en juego a actores tipo PYME y poniendo al alcance prácticamente de cualquier usuario disfrutar de todos estos servicios.

En el camino que hemos seguido hasta aquí también nos hemos encontrado con ciertas dificultades, tanto en el plano técnico como en el organizativo. En este sentido, el uso de la fuente abierta ha sido clave para la consecución de los objetivos, pues ha permitido personalizar los distintos componentes para ajustarlos a las necesidades del escenario. En el plano organizativo, las dificultades derivaron fundamentalmente de las diferencias en cuanto a la infraestructura técnica y a los objetivos de negocio de cada uno de los socios. A pesar de que los modelos de licencia disponibles en la comunidad de *software* libre son lo suficientemente flexibles como para compatibilizar las distintas estrategias empresariales, la gestión de la evolución de las diferentes piezas de la arquitectura ha supuesto un esfuerzo añadido.

La interacción entre los socios como miembros de la comunidad que conforma el proyecto se facilita por medio de la utilización de herramientas de colaboración como las listas de distribución, foros de discusión o los servidores de gestión de configuración.

Agradecimientos

Los autores desean agradecer a las empresas Telvent Interactiva y Telefónica I+D por su cooperación en la realización del trabajo aquí descrito y que forma parte de nuestra contribución al proyecto OSMOSE (Eureka 2023, ITEA ip00004), financiado parcialmente por la empresa Telvent y por el Ministerio de Ciencia y Tecnología, bajo referencia TIC2002-10373-E.

Referencias

[1] J. L. Ruiz, J. L. Arciniegas, R. Cerón, J. Bermejo, J. C. Dueñas. "A Framework for Resolution of Deployment Dependencies in Java-Enabled Service Gateways". Prooceedings on the FIDJI, Third International Workshop on Scientific Engineering of Distributed Java Applications, Springer Verlag LNCS, pg. 1-12, 2003.

[2] N. Serrano, S. Calzada, J.M. Sarriegui, I. Ciordia. "From proprietary to open source tools in information systems development". IEEE *Software*, pg. 59-66, 21, enero 2004.

[3] H. Kreger. "Java Management Extensions for application management". IBM Systems Journal, vol. 40, pg. 104-129, 2001.

[4] J. C. Dueñas, J. L. Ruiz, J. Bermejo, J. A. Alonso, C. Acuña, C. Díaz. "Plataformas abiertas para la provisión de servicios". Actas de las XIV Jornadas de Telecom I+D, noviembre 2004.

[5] M. A. Oltra, J. Bermejo, J. C. Dueñas, J. L. Arciniegas, C. Acuña, M. García. "Análisis de aspectos de seguridad en plataformas de servicios de gestión remota". Actas de las XIV Jornadas de Telecom I+D, noviembre 2004.

[6] J. L. Ruiz, J. C. Dueñas, F. Usero, C. Diaz. "Deployment in dynamic environments". Proceedings on DECOR, 1st francophone conference on deployment and (re)configuration, October 2004.

[7] OSGi Open Services Gateway Initiative, Web oficial disponible en: http://www.osgi.org.

[8] Telefónica I+D, Web oficial disponible en http://www.tid.es.

[9] Telvent Interactiva, Web oficial disponible en http://www.telvent.com.

[10] Objectweb Consortium, Web oficial disponible en http://www.objectweb.org.

[11] Apache *Software* Foundation, Web oficial disponible en http://www.apache.org.

[12] Proyecto *JBones*, Web oficial disponible en http://jbones.forge.os4os.org.

[13] Proyecto *JMood*, Web oficial disponible en http://jmood.forge.os4os.org.

[14] Proyecto *BarkIDS*, Web oficial disponible en http://barkids.forge.os4os.org.

[15] Proyecto *TIDuma*, Web oficial disponible en http://tiduma.forge.os4os.org

[16] Proyecto *TIDpema*, Web oficial disponible en http://tidpema.forge.os4os.org.

[17] Proyecto *DistrOS*, Web oficial disponible en http://distros.os4os.org

[18] Proyecto Oscar, Web oficial disponible en http://oscar.objectweb.org.

[19] OSS/J initiative, Web oficial disponible en htttp://www.ossj.com.

[20] R. S. Hall, H. Cervantes. “Challenges in Building Service-Oriented Applications for OSGi”. IEEE Communications, vol 42, pg 144-149, May 2004.

[21] Proyecto EUREKA ITEA Osmose, web oficial disponible en http://www.itea-osmose.org

Sistema maleable para el apoyo y guiado del aprendizaje colaborativo basado en servicios grid

Miguel L. Bote Lorenzo, Eduardo Gómez Sánchez, Guillermo Vega Gorgojo,
Yannis. A. Dimitriadis, Juan I. Asensio Pérez, Davinia Hernández Leo
Dpto. de Teoría de la Señal, Comunicaciones e Ingeniería Telemática. Universidad de Valladolid
ETSI de Telecomunicación. Camino Viejo del Cementerio s/n
47011 – Valladolid (Valladolid)
E-mail: {migbot, edugom, guiveg, yannis, juaase, davher}@tel.uva.es

***Abstract**. This paper introduces Gridcole, a system that can be easily tailored by educators in order to support the realization of collaborative learning scenarios designed by themselves. To do so, educators can provide a script specifying the sequence of learning activities to be performed by students as well as the documents and tools required to support them. Gridcole can then search for these tools in a service-based grid in order to make them available to users whenever it is required. Significantly, these tools are not limited in terms of access to supercomputational capabilities or specific hardware resources. Furthermore, Gridcole can guide students during the realization of the collaborative learning scenario according to the sequence of activities defined in the script.*

1 Introducción

El aprendizaje colaborativo [1] es un proceso en el que la adquisición de conocimientos y habilidades se lleva a cabo mediante la interacción entre los participantes del mismo. Esta aproximación pedagógica es, en muchas circunstancias, más efectiva que el aprendizaje individual [2], pero requiere un mayor esfuerzo por parte del educador para diseñar situaciones de aprendizaje colaborativo. Una situación de aprendizaje colaborativo se define como un escenario creado con la intención de que los alumnos construyan conocimiento a través de la realización de una serie de actividades de aprendizaje colaborativo y de aprendizaje individual [3].

En este sentido, el Aprendizaje Colaborativo Apoyado por Ordenador (CSCL – *Computer Supported Collaborative Learning*) [4] es una disciplina que estudia el uso de las tecnologías de información y comunicaciones (TIC) como herramientas de mediación para facilitar el aprendizaje colaborativo. Fruto de la intensa actividad en este dominio durante los últimos años, son numerosos los sistemas de aprendizaje colaborativo apoyado por ordenador (o sistemas CSCL) desarrollados hasta el momento. En general, un sistema CSCL es una aplicación que integra un conjunto de herramientas software de apoyo para la realización de una situación de aprendizaje colaborativo. Estas herramientas pueden ser específicas de un dominio de aprendizaje concreto (ej: un editor colaborativo de cadenas de energía para aprender Física) o genéricas (ej: un editor de textos). Además, las herramientas pueden ser tanto de uso individual como colaborativo.

Una de las características más deseadas en los sistemas de aprendizaje en general, y en los de aprendizaje colaborativo en particular, es la maleabilidad (*tailorability*). Un sistema maleable es aquel que permite a sus usuarios añadir nuevas funcionalidades. De esta manera, un sistema CSCL maleable permite modificar el conjunto de herramientas ofrecido para apoyar un escenario de aprendizaje colaborativo dado. Algunos ejemplos de este tipo de sistemas son DARE [5], Symba [6] y CURE [7]. Sin embargo, la utilidad de los sistemas de aprendizaje colaborativo maleables que es posible encontrar en la literatura se ve limitada en dos aspectos fundamentales.

Por una parte, no ofrecen la posibilidad de integrar herramientas que hagan uso de capacidades de supercomputación o de recursos de hardware específico. Existen numerosas situaciones de aprendizaje colaborativo en las que este tipo de herramientas es necesario, especialmente en áreas como las ciencias naturales, la ingeniería o la medicina. Por ejemplo, en COVASE [8] se utilizan recursos de supercomputación para crear un entorno de realidad virtual en el que los alumnos interaccionan con modelos tridimensionales complejos relacionados con el análisis de elementos finitos o la mecánica de fluidos. En PEARL [9] se accede a hardware específico (un generador de señales, un osciloscopio y un microcontrolador) como parte del proceso de aprendizaje colaborativo en el contexto de un laboratorio de electrónica. Sin embargo, estos sistemas no son maleables, y por lo tanto impiden al educador utilizarlos en situaciones de aprendizaje colaborativo distintas.

Por otra parte, los sistemas maleables existentes tampoco ofrecen la posibilidad de interpretar guiones colaborativos definidos por el educador. Un guión colaborativo es un conjunto de instrucciones que define, entre otras cosas, qué secuencia de actividades deben realizar los alumnos y cómo éstos

deben colaborar [10] para lograr unos objetivos pedagógicos determinados. De acuerdo con [10], la realización de actividades de aprendizaje de acuerdo con un guión colaborativo permite aumentar la efectividad del aprendizaje. La interpretación de un guión colaborativo por parte de un sistema CSCL implica que éste debe hacerse cargo de gestionar la secuencia de actividades que deben realizar los participantes del escenario. Un ejemplo de sistema CSCL guiado es Universanté [11]. Sin embargo, se trata de un sistema no maleable que no permite al educador modificar el guión utilizado

Este artículo propone Gridcole, un sistema que explota los beneficios de los servicios grid para permitir al educador integrar de manera sencilla herramientas que respondan a sus necesidades educativas, incluyendo herramientas con necesidades de supercomputación o de recursos de hardware específico. Además, Gridcole es un sistema dotado con la capacidad de interpretar guiones colaborativos basados en IMS-LD (*IMS Learning Design* – Diseño de Aprendizaje IMS) [12] definidos por el educador para, de este modo, permitir realizar actividades de aprendizaje colaborativo de forma guiada.

El resto del artículo se estructura de la siguiente manera. En la sección 2 se discuten las tecnologías empleadas para abordar los problemas detectados en los sistemas maleables. La sección 3 introduce un nuevo sistema que, empleando dichas tecnologías, pretende superar las limitaciones antes mencionadas. La sección 4 muestra dos ejemplos de situaciones de aprendizaje colaborativo que podrían ser apoyadas por este nuevo sistema. Finalmente, la sección 5 recoge las principales conclusiones de este trabajo.

2 Tecnologías usadas en Gridcole

La especificación IMS-LD y el grid basado en servicios pueden ser empleados para abordar las limitaciones de los sistemas maleables detectados en la sección anterior. Esta sección discute cómo pueden emplearse ambas tecnologías para construir un nuevo sistema que permita el guiado de situaciones de aprendizaje colaborativo en las que se pueda utilizar herramientas que hagan uso de recursos de supercomputación o de hardware específico.

2.1 IMS-LD para la formalización de guiones colaborativos

IMS-LD [12] es un lenguaje de modelado educativo basado en XML para la formalización de procesos de enseñanza-aprendizaje en documentos denominados diseños de aprendizaje. Concretamente, IMS-LD permite la descripción de *escenarios de aprendizaje* en términos de un *flujo de actividades* y un conjunto de *entornos*. El flujo de actividades especifica la secuencia de *actividades* que un aprendiz debe llevar a cabo para lograr determinados objetivos pedagógicos en función del rol que éste desempeña en el escenario. Los entornos se describen en términos de las *herramientas* y *contenidos* que se deben poner a disposición de los alumnos a la hora de realizar cada actividad. La popularidad de la especificación IMS-LD ha dado lugar a la aparición de herramientas de autoría como Reload [13] que permiten a los educadores generar diseños de una manera más sencilla.

IMS-LD suele ser utilizado en combinación con otras especificaciones. De esta manera, es posible emplear la especificación IMS-LRM (*IMS Learning Resource Metadada* – Metadatos para Recursos de Aprendizaje IMS) [14] para introducir en los diseños de aprendizaje descripciones genéricas de los documentos y las herramientas que se utilizan en el apoyo de las distintas actividades. Además, los diseños de aprendizaje suelen incluirse en un fichero denominado *unidad de aprendizaje* de acuerdo con la especificación IMS-CP (*IMS Content Packaging* – Empaquetado de Contenidos IMS) [15]. En las unidades de aprendizaje es posible incluir referencias a los documentos y herramientas concretos que deben ser empleados para apoyar el diseño empaquetado.

IMS-LD también puede ser empleado para la formalización de guiones colaborativos en combinación con la extensión propuesta en [16]. Esto permite que los guiones puedan ser interpretados de manera automática por motores de flujo de aprendizaje como Coppercore [17] que pueden ser incluidos en el contexto de un sistema de e-aprendizaje.

2.2 Servicios grid para el apoyo de actividades de aprendizaje

La computación grid es un paradigma computacional que promueve la compartición de todo tipo de recursos de software y hardware entre múltiples organizaciones administrativas [18]. Las aplicaciones que se ejecutan en un grid pueden hacer uso de los recursos ofrecidos en él por cualquier organización Dichos recursos pueden ser utilizados por las aplicaciones con cualquier propósito.

En particular, el grid computacional ha sido tradicionalmente empleado en aplicaciones de ámbito científico. Sin embargo, en la actualidad se considera que la educación será una de las aplicaciones del grid más importantes en un futuro cercano [19,20]. En esta subsección se discute cómo un grid puede ofrecer herramientas susceptibles de ser integradas en un sistema de e-aprendizaje maleable.

En un grid computacional, distintas organizaciones proveedoras podrían ofrecer todo tipo de herramientas potencialmente útiles para apoyar diferentes actividades de aprendizaje, tanto de tipo colaborativo como individual. De acuerdo con el marco conceptual definido por OGSA (*Open Grid Service Architecture* – Arquitectura Abierta de Servicios Grid) [18], dichas herramientas deberían ser expuestas como servicios grid. Además, estos

servicios podrían ser registrados en directorios específicos para herramientas de apoyo al aprendizaje. Esto facilitaría a los educadores localizar las herramientas que necesitasen para dar apoyo a su situación de aprendizaje colaborativo particular.

Estas herramientas deberían ser desarrolladas por los propios proveedores o, de forma alternativa, integradas total o parcialmente a partir de servicios grid compartidos por otros proveedores. En particular, las herramientas podrían hacer uso de manera agregada de los recursos computacionales disponibles en el grid para abordar problemas de supercomputación. Igualmente, podrían acceder a aquellos recursos hardware compartidos en el grid de los que el proveedor no disponga de forma local. Esto no implica, sin embargo, que las herramientas ofrecidas por los proveedores se reduzcan a aquellas que hacen uso de recursos extraordinarios compartidos por terceros. En cambio, esto significa que el grid ofrecería todo tipo de herramientas de apoyo al aprendizaje que, en aquellas situaciones en que sea necesario, no se verían limitadas en cuanto al uso de dichos recursos.

Teniendo en cuenta que la interacción con el usuario será en la mayor parte de las ocasiones un aspecto muy importante de las herramientas de apoyo al aprendizaje, los proveedores de dichas herramientas deberían proporcionar no sólo la implementación de la lógica de aplicación de las mismas de acuerdo con la especificación OGSI (*Open Grid Service Infrastructure* – Infraestructura Abierta de Servicios Grid) [21], sino también el cliente del servicio con la implementación de la lógica de presentación del mismo. En la actualidad OGSA no dispone de una especificación formal para los servicios orientados a presentación como es el caso de la especificación WSRP (*Web Services for Remote Portlets* – Servicios Web para Portlets Remotos) [22] para servicios web. Sin embargo, de acuerdo con las ideas de ésta, es razonable pensar que los clientes de los servicios grid se deberían ajustar a un estándar que facilitara, entre otros aspectos, la instalación del cliente en una máquina distinta a la del proveedor.

De este modo, las herramientas ofrecidas como servicios grid podrían ser fácilmente empleadas en el contexto de un sistema de aprendizaje colaborativo. Para ello bastaría con que el sistema ofreciera los distintos clientes de servicio que permiten a los usuarios utilizar las herramientas. En el caso concreto de los sistemas maleables, los educadores deberían poder añadir al sistema nuevos clientes de servicio para así permitir el uso de nuevas herramientas en la realización de una situación de aprendizaje dada. Esta aproximación a la maleabilidad se corresponde con el modelo de integración blanda definido en [23].

En la literatura es posible encontrar diversas propuestas [24,25] de sistemas de e-aprendizaje que utilizan un grid basado en servicios para poder acceder a capacidades de supercomputación o recursos hardware específicos. Sin embargo, ninguno de ellos es un sistema maleable. Tampoco ofrecen la posibilidad de guiado de los estudiantes a través de la interpretación de un guión.

3 El sistema Gridcole

Gridcole es un sistema maleable que puede ser empleado para apoyar situaciones de aprendizaje colaborativo guiado definidas con IMS-LD utilizando herramientas basadas en servicios grid. Esta sección describe su arquitectura y funcionamiento.

3.1 Usuarios definidos

En el sistema Gridcole es posible distinguir hasta un total de cuatro tipos de usuarios diferentes. Éstos se describen a continuación.

Los **aprendices** son aquellas personas que participan en la realización de situaciones de aprendizaje colaborativo apoyadas por el sistema para alcanzar determinados objetivos educativos.

Los **educadores** son los encargados de proporcionar al sistema las situaciones de aprendizaje que van a ser realizadas por los aprendices de acuerdo con los objetivos pedagógicos perseguidos. Los educadores también pueden participar en la realización de las situaciones de aprendizaje colaborativo.

Los **administradores** son las personas responsables de la gestión de diferentes aspectos del sistema tales como las operaciones de alta, baja y modificación de los distintos tipos de usuarios.

Los **proveedores** son aquellas organizaciones que ofrecen cualquier tipo de herramientas susceptibles de ser integradas en el sistema para el apoyo de actividades realizadas en el contexto de una situación de aprendizaje colaborativo. En Gridcole dichas herramientas pueden ser tanto aplicaciones autónomas (*stand-alone*) como herramientas basadas en servicios grid que, en caso de necesitarlo, hacen uso de los recursos de supercomputación y los recursos de hardware específico compartidos por otras organizaciones.

3.2 Funcionalidad básica

Gridcole es un sistema que permite a los educadores integrar las herramientas adecuadas para la realización de una situación de aprendizaje colaborativo dada. Para ello el educador debe proporcionar al sistema una unidad de aprendizaje que contengan la descripción formal basada en IMS-LD correspondiente. El sistema admite dos tipos de unidades: completas e incompletas.

Las unidades completas son aquellas en las que se incluye toda la información necesaria para que el sistema pueda saber qué herramientas en concreto van a ser utilizadas para el apoyo de las actividades

definidas en la situación de aprendizaje colaborativo de acuerdo con la especificación IMS-CP. En las unidades incompletas no se incluye esta información, sino simplemente una descripción genérica de cada una de las herramientas de acuerdo con la especificación IMS-LRM. En caso de que la unidad proporcionada al sistema sea de tipo incompleto, Gridcole se encarga de buscar entre las herramientas ofrecidas por los proveedores para seleccionar las que mejor se ajustan a dichas descripciones. De esta manera, el educador puede escoger aquellas que considere más adecuadas para el apoyo de la situación incluida en la unidad.

Gridcole permite que aprendices y educadores participen en la realización de situaciones de aprendizaje colaborativo. Durante esta realización, el sistema guía a los participantes indicándole a cada uno de ellos cuál es la actividad que le corresponde hacer de acuerdo con su rol y la secuencia definida en la situación de aprendizaje. Además, el sistema les ofrece la posibilidad de utilizar las herramientas y documentos que han sido especificados para el apoyo de la actividad indicada.

3.3 Arquitectura

La arquitectura del sistema Gridcole, tal y como se puede apreciar en el esquema que se muestra en la Fig. 1, consta de tres componentes principales que operan en el contexto de un grid computacional basado en servicios. Éstos son un portal web, un cliente y un motor de flujo de aprendizaje. Dichos componentes se describen a continuación.

El **portal web** proporciona a los usuarios un punto de acceso único al sistema. El portal es responsable de autenticar a los usuarios y de proporcionarles la asistencia adecuada en función de las distintas operaciones que puede realizar cada tipo de usuario en el sistema. Para ello el portal cuenta con una serie de elementos que es necesario mencionar.

Uno de estos elementos es una *base de datos* en la que el sistema mantiene información administrativa como el nombre, la clave y el perfil correspondiente a cada usuario. En la base de datos también se almacena información acerca de las unidades de aprendizaje que se encuentran en proceso de ejecución en cada momento así como de los usuarios que están autorizados para participar en las mismas. En el *repositorio de unidades*, en cambio, se almacenan todas las unidades de aprendizaje completas e incompletas que están disponibles para su uso en el sistema.

Otro elemento importante es *el buscador de herramientas*, el cual se encarga de encontrar las herramientas disponibles en el contexto del sistema que más se ajusten a las descripciones incluidas en una unidad de aprendizaje incompleta. Para ello, el buscador consulta los *registros de herramientas* en los que los distintos proveedores publican toda la información necesaria para uso y enlazado de dichas herramientas en el contexto de un sistema de aprendizaje junto con una descripción de las mismas basada en la especificación IMS-LRM. Finalmente, el *gestor de seguridad* es el elemento encargado de proporcionar a los usuarios del sistema los certificados temporales adecuados con los que es posible acceder a aquellas herramientas que son compartidas en el grid de manera segura.

El **motor de flujo de aprendizaje** es el componente del sistema que interpreta el guión colaborativo basado en IMS-LD correspondiente a cada unidad de aprendizaje que se encuentre en ejecución. De esta manera, el motor de flujo se encarga de determinar qué actividades debe realizar cada usuario que participa en la ejecución de la unidad, de qué herramientas y documentos puede disponer en cada actividad y cuándo ha de realizarla. También pone en contacto instancias concretas del cliente y el servicio de una herramienta dada.

El **cliente** proporciona una interfaz gráfica en la que, de acuerdo con la información obtenida del motor de flujo, se indica al usuario qué actividad le corresponde realizar y de qué documentos y herramientas dispone para ello. Si el usuario selecciona uno de los documentos, el cliente se encarga de descargarlo para así ponerlo a su disposición. En cambio, si el elemento seleccionado es una herramienta ofrecida por el proveedor como un servicio grid, lo que se descarga y lanza automáticamente es el software que implementa la lógica de presentación de dicho servicio (es decir, el cliente del propio servicio). De manera análoga, si lo que selecciona el usuario es una herramienta ofrecida por el proveedor como una aplicación autónoma, ésta es descargada y lanzada de manera automática.

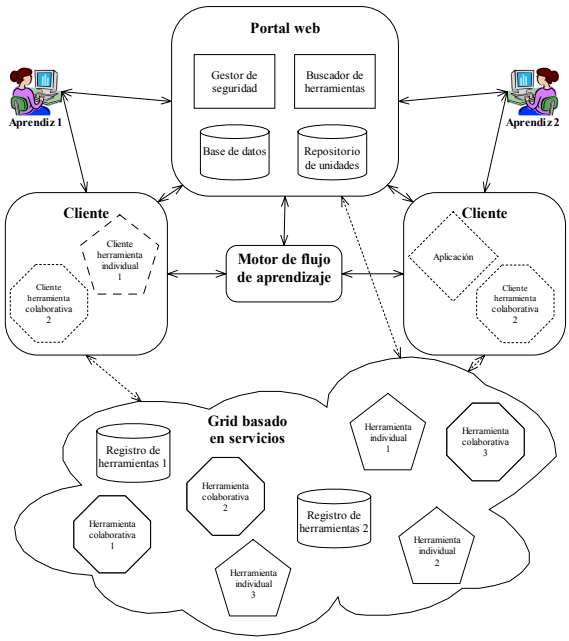


Figura 1: Arquitectura general del sistema maleable de aprendizaje colaborativo Gridcole.

3.4 Funcionamiento

Gridcole permite almacenar unidades de aprendizaje completas e incompletas, pero sólo puede ejecutar las primeras. Para facilitar la generación de una unidad completa a partir de una incompleta, el portal lee la descripción de las herramientas necesarias proporcionada mediante IMS-LRM, consulta los registros de herramientas ofrecidas por proveedores, y genera un listado de las que se adecuen a la descripción. Una vez seleccionada una de las herramientas por parte del educador, el portal le ofrece la posibilidad de configurar los parámetros de funcionamiento de la misma. Tras llevar a cabo esta misma operación con todas las herramientas descritas en el diseño de aprendizaje, el portal genera una nueva unidad completa en la que incluye toda la información de localización y configuración de las herramientas que han sido seleccionadas.

Cuando el educador decide ejecutar una unidad completa, el portal muestra la lista de usuarios del sistema así como los distintos roles definidos en la situación de aprendizaje. De este modo, el educador puede especificar qué usuarios van a participar en la realización de la situación y qué roles van a desempeñar cada uno de ellos. A continuación, el portal crea las instancias de los servicios grid no persistentes necesarios para el apoyo de la situación. Finalmente, el portal pasa la lista de usuarios, sus roles, y la localización de los servicios creados junto con la unidad de aprendizaje al motor de flujo para que éste se encargue de la ejecución de la misma.

Una vez comenzada la ejecución de la unidad, los usuarios seleccionados por el educador pueden participar en la misma. Así, el portal se encarga de crear un certificado temporal que permitirá a cada usuario emplear las herramientas basadas en servicios grid durante la ejecución de la unidad. A continuación, el portal pone al usuario en contacto con el cliente del sistema, el cual es configurado automáticamente para que pueda comunicarse con el motor de flujo.

Durante la ejecución de la unidad, el cliente se encarga de obtener del motor de flujo toda la información que debe mostrar al usuario en cada momento (ej: actividad que debe realizar, documentos y herramientas disponibles, etc.). Si el usuario decide utilizar una herramienta basada en servicios grid, el cliente se pone en contacto con el motor para averiguar todos los datos necesarios para su utilización: de dónde es posible descargar la lógica de presentación correspondiente, dónde se encuentra la instancia del servicio grid que se debe utilizar y cuáles son los parámetros de configuración. Esta información, junto con la localización del certificado temporal del usuario, es empleada por el cliente del sistema para lanzar en la máquina del usuario un cliente específico de la herramienta seleccionada configurado correctamente.

3.5 Prototipo

El sistema Gridcole se encuentra actualmente en proceso de desarrollo. Sin embargo, se ha desarrollado un prototipo con funciones limitadas para probar la viabilidad de las ideas presentadas en este artículo. Este prototipo se basa en el cliente y el motor de flujo proporcionados por el proyecto Coppercore [17]. Ambos han sido modificados convenientemente para permitir a los usuarios utilizar herramientas individuales y colaborativas basadas en servicios grid durante la realización de una situación de aprendizaje colaborativo. Para ello es imprescindible que los clientes de estas herramientas sean ofrecidas por los proveedores de acuerdo con el modelo de distribución de aplicaciones basado en Java Web Start que ya se utiliza en otros sistemas basados en servicios grid [26]. Este prototipo, sin embargo, aún no ofrece las funcionalidades correspondientes al portal del sistema.

4 Ejemplos de uso de Gridcole

Gridcole permite al educador poner en práctica situaciones de aprendizaje colaborativo apoyado por ordenador adecuadas a sus necesidades. Esta sección describe dos ejemplos, para Arquitectura de Ordenadores y Teletráfico y Gestión, ambas asignaturas troncales de la titulación Ingeniero de Telecomunicación de la Universidad de Valladolid.

4.1 Arquitectura de Ordenadores

Arquitectura de Ordenadores es una asignatura que tiene como objetivo lograr que los alumnos entiendan y apliquen los principios básicos de diseño y evaluación de sistemas informáticos. Por este motivo se ha diseñado una situación para esta asignatura en la que los alumnos interpretan el papel de consultores que deben colaborar para decidir cuál es la máquina más adecuada a las necesidades de un cliente ficticio (ej: un hospital, un centro meteorológico), de entre una serie de máquinas reales.

La situación diseñada consta de una primera fase individual en la que los alumnos estudian un documento de requisitos y deben proponer un modelo de carga para su cliente. También individualmente estudian la documentación de las distintas máquinas y *benchmarks* disponibles, y proponen un plan experimental de evaluación. Luego, utilizando una de las herramientas proporcionadas por Gridcole, ejecutan los *benchmarks* sobre dichas máquinas y recopilan los resultados, proponiendo una recomendación a título individual de la mejor máquina para el cliente. A continuación se reúnen en parejas y discuten los documentos de solución generados por cada miembro, repitiendo los mismos pasos de forma colaborativa, hasta proponer una solución común. Para ello pueden utilizar una herramienta de discusión síncrona como un *chat*. A continuación pueden juntarse dos parejas y alcanzar una solución común, y así sucesivamente hasta que

todo el grupo acuerda una solución definitiva. La secuencia de actividades de esta situación se corresponde con la estructura colaborativa denominada pirámide.

La realización de esta situación en un sistema maleable de aprendizaje de colaborativo implica que éste ofrezca dos características. Por un lado, la posibilidad de integrar una herramienta que hace uso de hardware específico. Éste es el caso de la herramienta de *benchmarking*, que debe permitir la posibilidad de ejecutar distintos *benchmarks* en un conjunto de máquinas con distintas arquitecturas que puedan ser consideradas de interés por los responsables de la asignatura. Por otro, la posibilidad de ejecutar guiones colaborativos. La situación considerada implica la realización de una secuencia de actividades muy precisa definida por la estructura colaborativa de pirámide. Para que los alumnos se puedan aprovechar el beneficio que supone la puesta en práctica de este tipo de estructuras es necesario que el sistema de aprendizaje no sólo facilite su seguimiento, sino que además induzca a los estudiantes a hacerlo. Ambas características están presentes en Gridcole.

Esta afirmación se ve corroborada por el hecho de que el prototipo de Gridcole ha sido empleado para realizar esta situación con alumnos ficticios bajo condiciones de laboratorio. Para el apoyo de dicha realización se han empleado sendas herramientas de *benchmarking* y *chat* desarrolladas por los autores de este trabajo y ofrecidas como servicios grid. En la Fig. 2 es posible observar un extracto del diseño de esta situación mientras que en la Fig. 3 se muestra una captura de pantalla del prototipo de Gridcole durante la realización de la misma.

4.2 Teletráfico y Gestión

El principal objetivo de Teletráfico y Gestión es completar la visión general de las redes y servicios de telecomunicación adquirida en otras asignaturas del área de Ingeniería Telemática mediante la puesta en práctica de muchos de los conceptos que ya han sido tratados en las mismas. De acuerdo con esto, se ha diseñado para esta asignatura una situación en la que los alumnos colaboran en el estudio de varios mecanismos de TCP (ej: ventana deslizante, inicio lento, evitación de la congestión, etc.).

Para la realización de esta situación los alumnos deben organizarse en grupos de cuatro, siendo cada uno de ellos *experto* en un mecanismo diferente. Durante la primera fase de la situación, de carácter individual, cada alumno debe estudiar el comportamiento del mecanismo asignado en un conjunto de escenarios seleccionados por el profesor

```
<manifest ...>
 <organizations>
  <imsld:learning-design identifier="LD-LAO" level="B" uri="">
  ...
   <imsld:learning-activity identifier="LA-realizar-benchmarks">
    <imsld:title>Realizacion de benchmarks</imsld:title>
    <imsld:environment-ref ref="E-realizar-benchmarks"/>
    ...
   </imsld:learning-activity>
   ...
   <imsld:environment identifier="E-realizar-benchmarks">
    <imsld:title>Entorno de benchmarking</imsld:title>
    <imsld:learning-object identifier="LO-benchmarking-tool">
     <imsld:title>Herramienta de benchmarking</imsld:title>
     <imsld:item identifier="I-tool-1" identifierref="RES-tool-1"/>
    </imsld:learning-object>
   </imsld:environment>
 ...
</imsld:learning-design>
 </organizations>
 <resources>
  <resource identifier="RES-tool-1" type="gridserviceclient"
href="http://egeo.tel.uva.es/clients/benchmarkingtoolclient.jnlp">
   <dependency  identifierref="RES-factory-1"/>
  </resource>
  <resource identifier="RES-factory-1" type="gridservicefactory"
href="http://egeo.tel.uva.es:8080/ogsa/services/BenchServFact"/>
 ...
 </resources>
</manifest>
```

Figura 2: Extracto del diseño en el que se define la actividad de benchmarking y se especifica la herramienta que se debe emplear.

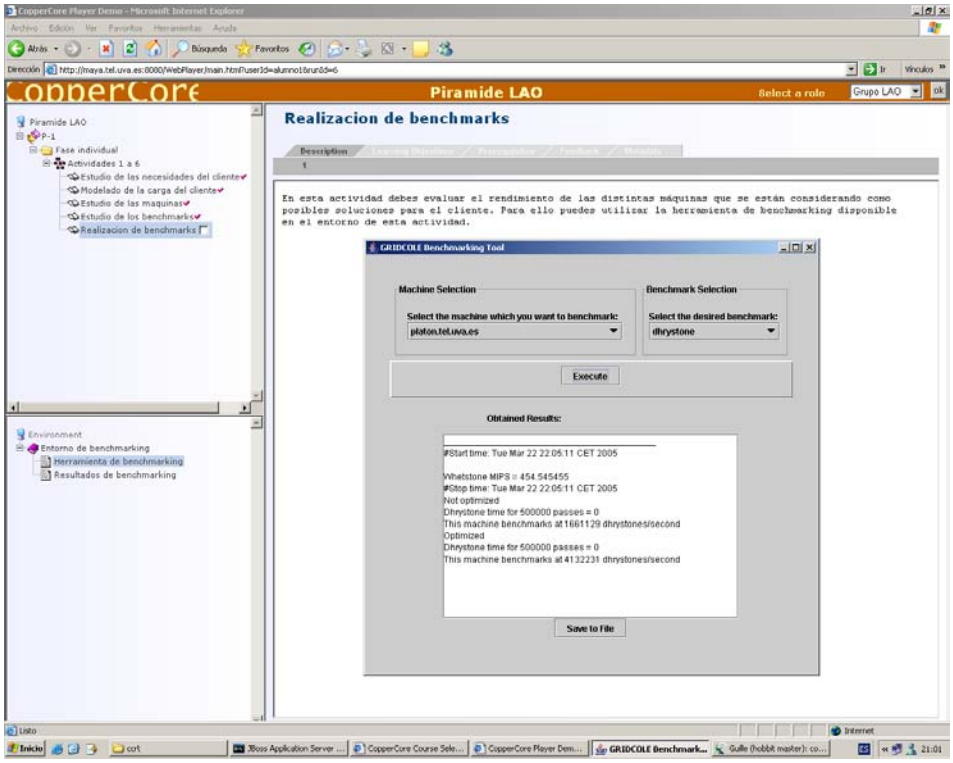


Figura 3: El cliente de Coppercore y la herramienta de *becnhmarking* durante la realización de la actividad de *benchmarking.*

para un amplio rango de parámetros (ej: probabilidades de pérdida, capacidades de enlace) del mismo utilizando una herramienta de simulación adecuada. Después, cada experto discute con los expertos del mismo mecanismo de otros grupos sobre las conclusiones a las que ha llegado utilizando una herramienta de discusión síncrona. Finalmente, los miembros de cada grupo discuten sobre la influencia que unos mecanismos tienen sobre otros que pueden observar en distintos escenarios de simulación también seleccionados por el profesor. La secuencia de actividades de esta situación se corresponde con la estructura colaborativa denominada rompecabezas.

La realización de este escenario en un sistema maleable también implica que éste muestre dos características. Por una parte, la posibilidad de integrar una herramienta que hace uso de capacidades de supercomputación. Este es el caso de la herramienta de simulación que necesitaría utilizar numerosos recursos de computación para poder concluir el barrido de parámetros en un plazo de tiempo razonable. Por otra, la posibilidad de ejecutar guiones colaborativos. Una vez más esta característica es necesaria para que el sistema pueda guiar a los alumnos a través de la secuencia de actividades definida para esta situación.

Gridcole también podrá ser utilizado para apoyar la realización de esta situación una vez que esté disponible la herramienta de simulaciones GIPSE [27]. Ésta es una herramienta basada en servicios grid que permite la realización de simulaciones con barrido de parámetros para el simulador de redes *ns-2* [28]. Mientras tanto, el prototipo ha sido empleado para comprobar que el sistema puede guiar a los alumnos a través de la secuencia de actividades definida para esta situación ofreciéndoles todos los documentos y herramientas necesarias para su realización excepto la herramienta de simulación.

5 Conclusiones y trabajo futuro

Gridcole es un sistema que permite a los educadores definir qué herramientas son necesarias para apoyar una situación de aprendizaje dada. A partir de esta información, Gridcole es capaz de enlazar dichas herramientas de acuerdo con el modelo de maleabilidad por integración blanda. Estas herramientas se ponen a disposición de los usuarios del sistema cuando son necesitadas.

A diferencia de otros sistemas maleables, Gridcole permite integrar herramientas ofrecidas por proveedores en el contexto de un grid computacional. Esto hace posible que los usuarios de Gridcole puedan emplear herramientas que necesiten hacer uso de capacidades de supercomputación o de recursos de hardware específico. Sin embargo, es importante aclarar que Gridcole no sólo permite utilizar este tipo de herramientas, sino que también es posible ofrecer como servicios grid herramientas que no hacen uso de capacidades de supercomputación o de recursos de hardware específico. Adicionalmente, estas herramientas "normales" también pueden ser integradas en Gridcole como aplicaciones autónomas.

Además, Gridcole es un sistema capaz de ejecutar guiones colaborativos formalizados empleando la especificación IMS-LD. En dicha ejecución el sistema se encarga de guiar a los alumnos indicándoles la secuencia de actividades que deben realizar y poniendo a su disposición los documentos y las herramientas que han sido definidos para su apoyo. En este sentido, es importante mencionar que el uso de la característica de guiado no es obligatorio. Si un educador desea utilizar Gridcole para prestar apoyo a una situación de aprendizaje en la que los alumnos deban colaborar libremente, basta con que cree un diseño de aprendizaje basado en IMS-LD en el que se defina una única actividad de aprendizaje junto con todas las herramientas y documentos necesarios para su apoyo.

En lo que al trabajo futuro se refiere, está previsto que durante la primavera de 2005 se evalúe desde un punto de vista educativo el apoyo que Gridcole puede prestar para la realización de la situación de Arquitectura de Ordenadores en un contexto real. También está previsto que continúe el trabajo de desarrollo de Gridcole para disponer de un sistema con funcionalidad completa. Mientras tanto, ya se está investigando cómo mejorar algunos elementos de su arquitectura, como el buscador de herramientas. En este caso concreto se está investigando el uso de ontologías con el objetivo de mejorar la descripción y búsqueda de herramientas en el grid.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el proyecto del Ministerio de Ciencia y Tecnología español TIC-2002-04258-C03-02, la red de excelencia europea IST-FP6-507838 y el proyecto de la Comisión Europea EAC/61/03/GR009. Los autores también agradecen la colaboración en este trabajo al resto de miembros del grupo GSIC/EMIC.

Referencias

[1] P. Dillenbourg. *Collaborative Learning: cognitive and computational approaches*, Oxford, UK: Elsevier Science (1999).

[2] R. T. Johnson, D. W. Johnson. "An overview of cooperative learning". In: *Creativity and collaborative learning: a practical guide to empowering students and teachers*, eds. J. S. Thousand, A. Villa, A. Nevin. Baltimore, MD, USA: Brookes Press, pp. 31-44 (1994).

[3] C. A. Osuna Gómez, Delfos: un marco telemático educativo basado en niveles orientado a situaciones de aprendizaje cooperativo 2000. Tesis Doctoral. Universidad de Valladolid, España.

[4] T. Koschmann. *CSCL: theory and practice of an emerging paradigm*, Malwah, NJ, USA: Lawrence Erlbaum (1996).

[5] G. Bourguin, A. Derycke. "Integrating the CSCL activities into virtual campuses: foundations of a new infrastructure for distributed collective activities". Proc. of the European Conf. on CSCL (Euro-CSCL 2001). Maastricht, The Netherlands, 2001.

[6] M. L. Betbeder, P. Tchounikine. "Symba, a tailorable framework to support collective activities in a learning context". Proc. of the 9th Int. Workshop on Groupware (CRIWG 2003). Grenoble, France, 2003.

[7] J. M. Haake, T. Schümmer, A. Haake, M. Bourimi, B. Landgraf. "Two-level tailoring support for CSCL". Proc. of the 9th Int. Workshop on Groupware (CRIWG 2003). Grenoble, France, 2003.

[8] N. Jensen, S. Seipel, W. Nejdl, S. Olbrich. "COVASE: Collaborative visualization for constructivist learning". Proc. of the Conf. on CSCL (CSCL 2003). Bergen, Norway, 2003.

[9] J. M. Martins Ferreira, G. R. C. Alves, R. Costa, N. Hine. "Collaborative learning in a web-accesible workbench". Proc. of the 8th Int. Workshop on Groupware (CRIWG 2002). La Serena, Chile, 2002.

[10] P. Dillenbourg. "Over-scripting CSCL: the risks of blending collaborative learning with instructional design". In: *Three worlds of CSCL. Can we support CSCL*, ed. P. A. Kirschner. Heerlen, Open Universiteit Nederland, pp. 61-91 (2002).

[11] A. Berger, R. Moretti, P. Chastonay, P. Dillenbourg, A. Bchir, R. Baddoura, C. Bengondo, D. Scherly, P. Ndumbe, P. Farah, B. Kayser. "Teaching community health by exploiting Int. socio-cultural and economical differences". Proc. of the European Conf. on CSCL (Euro-CSCL 2001). Maastricht, The Netherlands, 2001.

[12] IMS Learning Design Information Model specification v1.0: http://www.imsproject.org/.

[13] Reload Project: http://www.reload.ac.uk.

[14] IMS Learning Resource Metadata specification v1.2: http://www.imsglobal.org/.

[15] IMS Content Packaging Information Model specification v1.1.3: http://www.imsglobal.org/.

[16] D. Hernández-Leo, J. I. Asensio-Pérez, Y. Dimitriadis. "IMS Learning Design Support for the Formalization of Collaborative Learning Patterns". Proc. of the 4th Int. Conf. on Advanced Learning Technologies (ICALT 2004). Joensuu, Finland, 2004.

[17] CopperCore Project: http://coppercore.org/.

[18] I. Foster, C. Kesselman, S. Tuecke. "The Open Grid Services Architecture". In: *The Grid 2: blueprint for a future computing infrastructure*, eds. I. Foster, C. Kesselman. San Francisco, CA, USA: MK Publishers (2004).

[19] G. Fox. "Education and the enterprise with the grid". In: *Grid computing: making the global infrastructure a reality*, eds. F. Berman, G. Fox, A. Hey. Chichester, UK: John Wiley & Sons, pp. 963-976 (2003).

[20] L. Smarr. "Grids in context". In: *The Grid 2: blueprint for a future computing infrastructure*, eds. I. Foster, C. Kesselman. San Francisco, CA, USA: MK Publishers (2004).

[21] Open Grid Services Infrastructure specification v1.0: http://www.ggf.org.

[22] Web Services for Remote Portlets specification: http://www.oasis-open.org.

[23] A. Morch. "Three levels of end-user tailoring: customization, integration and extension". Proc. of the 3rd Decennial Aarhus Conf.. Aarhus, Denmark, 1995.

[24] A. Bagnasco, A. M. Scappola. "A grid of remote laboratory for teaching electronics". Proc. of the 2nd Int. workshop on e-Learning and Grid technologies. Paris, France, 2003.

[25] S. Wesner, K. Wulf, M. Müller. "How grid could improve e-learning in the environmental science domain". Proc. of the 1st Int. workshop on e-Learning and Grid technologies. Laussane, Switzerland, 2002.

[26] S. Peltier, M. E. Martone, S. Lamont, A. Lin, D. Lee, T. Molina, L. Dai, M. Wong, S. Mock, M. H. Ellisman. "The Telescience portal for advanced tomography applications". Journal of parallel and distributed computing, pp. 539-550, vol. 63 (5) (2003).

[27] J. M. Wozniak, A. Striegel, D. Salyers, J. A. Izaguirre. "GIPSE: streamlining the management of simulation on the grid". Proc. of the 38th Annual Simulation Symposium (ANSS'05). San Diego, CA, USA, 2005.

[28] The Network Simulator - ns-2: http://www.isi.edu/nsnam/ns/.

Diseño Eficiente de Aplicaciones Multimedia sobre Redes Inalámbricas Aplicando Programación Orientada a Aspectos

Miguel A. Quintana Suárez, Salvador Galván Sánchez, Elsa Mª Macías López, Alvaro Suárez Sarmiento

GAC (Grupo de Arquitectura y Concurrencia)
Departamento de Ingeniería Telemática. Universidad de Las Palmas de G. C.
Campus Universitario de Tafira
35017 Las Palmas de G.C.

mquintana@dit.ulpgc.es, salvador@galvansanchez.com, {emacias, asuarez}@dit.ulpgc.es

***Abstract**. Nowadays, it is very important the development of applications for wireless networks and mobile devices. However, the traditional programming methodologies for these devices are not appropriated because to the adaptation of these applications is a very tedious task. In this paper we propose the Aspect Oriented Programming (AOP) to develop multimedia applications over wireless networks and mobile devices. The different modules to be used by the applications will be woven with the necessary aspects. We have focused on detecting the wireless channel state to retrieve useful information that can be harnessed by the applications. With this approach we pretend to support QoS for these applications over wireless networks.*

1 Introducción

En los últimos años hemos asistido a una gran evolución de las redes inalámbricas y los servicios ofertados sobre ellas. Entre las diferentes tecnologías y estándares, actualmente el estándar IEEE 802.11 es el más utilizado. Son muchos los trabajos desarrollados en aras de soportar QoS para aplicaciones multimedia [1] sobre estas redes.

En la actualidad, entre los diferentes tipos de servicios multimedia, el servicio de media *streaming* [2] es uno de los más demandados. Por otra parte, las aplicaciones cooperativas existentes están basadas en el uso de elementos como video streaming, visualización compartida de documentos, chats, mensajería y permiten cooperar a los usuarios de manera distribuida. En los últimos tiempos se han desarrollado varios de estos entornos para aplicaciones multimedia cooperativas y distribuidas [3]. Unos definen e implementan completamente la aplicación multimedia distribuida desde cero [4], esta técnica adolece de largos tiempos de desarrollo y poca adaptabilidad a cambios para introducir mejoras del servicio. Otros utilizan una técnica que consiste en desarrollar la funcionalidad principal como una aplicación monousuario y luego transformarla en su versión distribuida, considerando su uso por varios usuarios. Un ejemplo es la edición cooperativa: primero se implanta un editor y después se añaden las herramientas de edición cooperativa y distribuida, añadiendo nuevos botones u opciones de menú. Esta técnica tiene el problema de que no puede ser integrada como un módulo separado con otras aplicaciones y la interfaz de usuario llega a ser difícil de utilizar. Otra posibilidad consiste en el desarrollo de un *middleware* que será utilizado para adaptar su ejecución, transformándolas en aplicaciones cooperativas distribuidas [5], pero esta técnica no es apropiada para adaptar la ejecución de aplicaciones a diferentes redes o dispositivos de naturaleza heterogénea.

Algunos de los problemas relacionados con estos dispositivos se deben a las diferencias en capacidad de memoria, resolución de pantalla o potencia de cálculo. Para este tipo de dispositivos existen trabajos que consideran la *Programación Orientada a Aspectos*, *AOP* [6], como una de las metodologías más adecuadas, aunque son pocos los que desarrollan aplicaciones para dispositivos móviles con acceso a redes inalámbricas [7]. En [8] este paradigma es utilizado a nivel de componentes y en [9] demuestran las ventajas de la refactorización del middleware usando AOP. Sin embargo, uno de los desafíos es la generación de aspectos que permitan implementar un middleware específico [10] y adaptarlo a dispositivos heterogéneos de terminales móviles y acceso inalámbrico.

En este trabajo presentamos nuestros resultados preliminares usando la AOP para transformar rápidamente una aplicación multimedia para que contemple las particularidades de funcionamiento de las redes inalámbricas. En tiempo de compilación se tejen junto a la aplicación los elementos necesarios para realizar las operaciones de: búsqueda de recursos, sincronización y reparto de contenidos multimedia, anticipación de estados indeseados en redes inalámbricas o detección de la disponibilidad del canal inalámbrico para evitar los perjuicios que estos pueden ocasionar sobre la ejecución de la aplicación.

El resto de este artículo está organizado como sigue: en la sección 2 revisamos algunas cuestiones relacionadas con el comportamiento de las redes inalámbricas en presencia de fallos en el canal cuando ejecutan aplicaciones. En el apartado 3 presentamos la AOP junto a la forma que utilizamos para aislar los aspectos; en el apartado 4 ayudados de una aplicación multimedia de reparto y visualización de vídeo, introducimos aspectos que pueden ser utilizados en el desarrollo de aplicaciones multimedia cooperativas y en la sección 5 aquellos otros elementos relacionados con redes inalámbricas. Para finalizar incluimos las referencias a otros trabajos realizados en la refactorización de middleware con AOP y para concluir destacamos cuales son algunas de las conclusiones obtenidas y las líneas de trabajo abiertas.

2 Identificación de Problemas en Redes Inalámbricas

La cobertura de un enlace radio se mide con el nivel de señal, el nivel de ruido y la relación señal a ruido (SNR). Existen herramientas de monitorización de la cobertura que presentan valores gráficamente [11]. Esta y otras herramientas similares interrogan al driver de la tarjeta inalámbrica para obtener esta información: otras informaciones disponibles podrían ser el número de puntos de acceso, el identificador de la red (*SSID*), el número de canales, la distancia a la que se encuentran los puntos de acceso, etc. Toda esta información se puede obtener desde llamadas a bibliotecas como NDIS [12], consultando un archivo en el sistema operativo Linux (*/proc/wireless*), desde programas en Java que importen el paquete definido por Place Lab [13], etc.

En cuanto a la congestión en el canal radio está demostrado que el protocolo TCP no se comporta bien [14] para las redes de área local que siguen el estándar IEEE 802.11 [15]. Tampoco existe una forma estándar y aceptada para informar al TCP desde el driver del nivel de enlace sobre desconexiones en el canal radio (problemas de cobertura) y pérdidas de paquetes por congestión en el canal.

En la práctica: 1) no existe ninguna herramienta que informe a los usuarios, aplicaciones, middlewares o protocolos que se ha producido una situación de falta de cobertura y la diferencie de una de excesiva congestión. 2) Si existiera esta herramienta sería muy compleja porque la acción que haría cada usuario, aplicación, middleware o protocolo no sería la misma sino que tendría que ser particularizada según el tipo de receptor de la información, el tipo de flujo que se pone en la red, etc.

En [16] se presenta una herramienta que es capaz de detectar si un dispositivo con conexión inalámbrica IEEE 802.11 entra en (o sale de) una zona donde la calidad de la señal inalámbrica se degrada por diversos motivos: presencia de obstáculos (p.e., paredes), ruido generado por otros dispositivos, condiciones climáticas, fallos en el enlace, congestión de la red, etc. La herramienta, que introduce baja sobrecarga en la red, se implanta sobre el dispositivo que monitoriza la cobertura y la calidad del enlace inalámbrico, midiendo parámetros como: 1) el tiempo de ida y vuelta (*RTT - Round Trip Time*) entre el dispositivo y el punto de acceso al que está asociado para lo cual utiliza las bibliotecas libpcap y libnet para capturar e inyectar paquetes *Internet Control Message Protocol* (*ICMP*) en la red; 2) el nivel de señal, de ruido y la relación señal a ruido, leyendo del archivo */proc/wireless*. Los autores demuestran que la mejor estimación de la cobertura la obtienen cuando combinan de forma apropiada los parámetros medidos. En caso contrario, la probabilidad de falsas alarmas aumenta. Esta herramienta puede ser utilizada de diferentes formas: por ejemplo, avisar al usuario para que se mueva hacia una zona donde la cobertura sea mejor, avisar al *driver* de la tarjeta inalámbrica para que pase a modo de ahorro de batería, posponer acciones críticas como la actualización de bases de datos, etc. Nosotros hemos utilizado esta herramienta para prevenir que las aplicaciones paralelas distribuidas síncronas queden bloqueadas a la espera de datos que no van a llegar porque se encuentran almacenados en dispositivos sin conexión hacia el proceso que recolecta esta información [17] [18].

3 Programación Orientada a Aspectos

La AOP es una nueva metodología de programación que permite a los desarrolladores agrupar en módulos independientes los distintos objetivos, funcionalidades, aspectos o intereses, llamados *concerns*, que aparecen diseminados sobre los distintos componentes que conforman el sistema. Esta técnica puede ser utilizada tanto para transformar una aplicación existente como para desarrollar nuevas aplicaciones desde el principio. En cualquier caso el estudio y separación de los diferentes *concerns* de una aplicación nos llevan a planificar el desarrollo del software desde un punto de vista diferente al tradicional. El beneficio obtenido por la orientación a aspectos se atribuye a la posibilidad de manejar y utilizar un paradigma de descomposición vertical con algunos de estos objetivos, debido a su transversalidad en el sistema.

La AOP define una serie de pasos que deben tenerse en cuenta a la hora de resolver un problema con esta metodología [19]:

- Descomposición en Aspectos: descomponer los requisitos del sistema para identificar cada una de las funcionalidades.

- Implantación de funcionalidades: se debe diseñar y desarrollar cada una ellas de manera independiente al resto.

- Recomposición de Aspectos: especificar las reglas que deben seguirse para (re)combinar los diferentes aspectos hasta obtener las unidades modulares del sistema.

La AOP vence las limitaciones de los paradigmas de programación tradicionales mediante la definición de un lenguaje que permita crear diferentes módulos para generar propiedades o funcionalidades, pudiendo éstas ser desarrolladas por separado de la funcionalidad principal. En particular, aquellas asociadas a requisitos extra-funcionales y que pueden encontrarse diseminadas a lo largo de la aplicación. El resultado final se genera a través de un proceso de tejido de aspectos (*aspect weaver*), que mezcla los módulos que contienen los aspectos y los que poseen las funcionalidades principales de la aplicación. Esta metodología dispone de entornos y lenguajes de programación concretos, como por ejemplo AspectJ [20]. AspectJ define un conjunto de elementos que dan a Java las características que permiten modelar aspectos. Estos son: puntos de enlace (*joinpoint*) que representa un punto de intercepción en el flujo normal de ejecución de la aplicación y puntos de corte (*pointcut*) que son usados para definir un conjunto de puntos de enlace. Antes, después o en vez de la ejecución de un punto de corte pueden ser ejecutadas las acciones concretas que están definidas en el aspecto: estas formas son conocidas como *advice*.

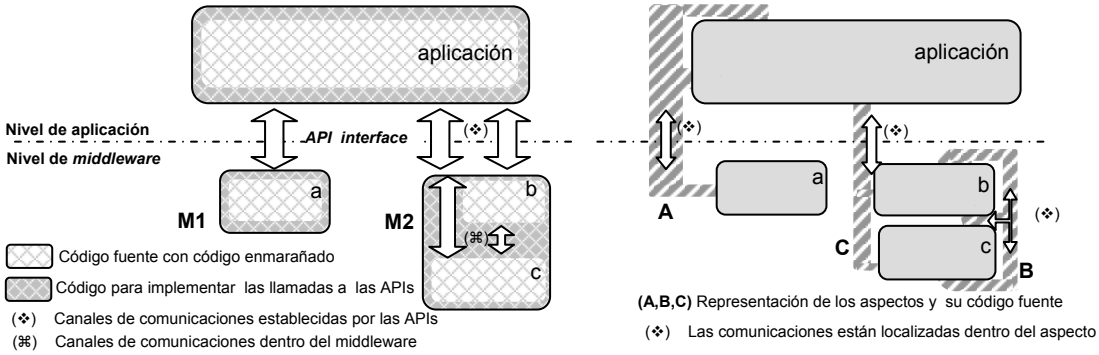
3.1 Aislando aspectos

En este apartado se presentan algunas ideas sobre el proceso de extracción, identificación e implantación de requisitos sean estos funcionales o no, *concerns*, candidatos a ser modelados como aspectos y cómo éstos pueden ser extraídos de las aplicaciones y middleware ya existentes. Para ilustrar el proceso de extracción e identificación de estos *concerns* utilizamos como ejemplo una aplicación que utiliza los servicios suministrados por un middleware.

Mediante la refactorización aislamos cada uno de los diferentes *concerns* existente en la aplicación o en el middleware. Estos *concerns* serán desarrollados de manera autónoma en aspectos independientes y posteriormente tejidos para crear nuevamente la aplicación.

En la figura 1.a representamos el enfoque tradicional de implantación de una aplicación que utiliza las facilidades suministradas por dos componentes middleware, llamados M1 y M2. Estos servicios están soportados sobre APIs, *Application Programming Interfaces*. Las llamadas a estas APIs pueden encontrarse a lo largo del código fuente de la aplicación y definen los diferentes canales de comunicación desde la aplicación hacia el middleware. No existen zonas restringidas dentro del código fuente de la aplicación donde puedan invocarse dichas llamadas. El uso de las primitivas definidas en el middleware está entremezclado, disperso y tejido con la implementación de la funcionalidad principal de la aplicación, definiendo múltiples puntos de enlace entre la aplicación y el middleware, a lo que denominamos canales de comunicación. Esa misma situación puede encontrarse dentro del propio middleware, como es el caso de M2. En la figura 1.a hemos representado ese código enmarañado como el mallado que recubre los elementos que implementan las distintas funcionalidades (objetos a, b, c y aplicación). La parte más oscura representa el código adicional necesario para permitir la interoperatividad entre ellos.

En la figura 1.b representamos un nuevo modelo de aplicación, refactorizada mediante el uso de la AOP. El procedimiento seguido se descompone en dos grandes pasos: en el primero, localizamos los diferentes elementos funcionales del middleware, los aislamos y los transformamos desde una implantación de llamadas a API a objetos individuales y por otro lado, identificamos las comunicaciones de estos con el resto de elementos. Ahora los objetos, (aplicación, a, b y c), representan elementos autónomos no diseñados expresamente para trabajar conjuntamente. Los elementos A, B y C, representan aspectos o *concerns* independientes. Dentro de cada uno de estos aspectos se define cual es la política a utilizar en las comunicaciones entre los distintos elementos que los



a) Utilización de servicios de un middleware en forma de API

b) Aplicación diseñada mediante el uso de la AOP

Figura 1. Dos esquemas de una misma aplicación

componen. La parte de los aspectos que está fuera de cualquiera de las cajas representa lo necesario, código propio del aspecto, para la interoperatividad conjunta de los elementos relacionados en el aspecto. La parte que aparece difuminada dentro de los rectángulos representa el código, que perteneciente al aspecto, debe ser insertado por el tejedor con el fin de atrapar diferentes eventos, inhibir la ejecución de algún método o simplemente hacer un traspaso de parámetros y definir los canales de comunicación necesarios. El motivo de que esté ligeramente difuminado es que su inserción es transparente a los objetos, ya sea a la aplicación o a cualquiera del resto de objetos.

Hemos resumido aquí uno de los puntos, que a nuestro entender, es de los más importantes en el proceso de utilización de la orientación a aspectos y que permite utilizar el middleware de manera distinta a la tradicional. Las ventajas de este modelo es que la modificación de cualquiera de los elementos presentes dentro de un aspecto por otro, que soporte la misma funcionalidad u otra equivalente, es muy sencillo al requerir cambiar únicamente el aspecto que los une y sobre todo, que el flujo de intercambio está perfectamente localizado al aspecto que le afecta.

4 Ejemplo de Aplicación

Teniendo en cuenta los problemas con las redes inalámbricas y los dispositivos móviles, y conociendo el modo en el que utilizamos los aspectos en el desarrollo de aplicaciones, diseñamos una aplicación donde podemos demostrar los beneficios obtenidos con esta nueva metodología. La aplicación multimedia desarrollada se compone de un cliente y un servidor para la visualización y transmisión de vídeo. La aplicación cliente muestra cada uno de los fotogramas de vídeo recibidos. Estos fotogramas son proporcionados por el servidor a través de la red.

La aplicación cliente, figura 2.a, tiene una interfaz gráfica compuesta por un visor de fotogramas de vídeo y por un grupo de botones que controlan el reproductor de medios, estos botones son el SETUP, PLAY, PAUSE y STOP. El servidor, figura 2.b, posee una interfaz gráfica que muestra los clientes que están conectados. La interfaz posee dos botones donde uno de ellos actualiza la lista de clientes actuales conectados al servidor y el otro está asociado a la función de arrancar y parar el servidor.

a) cliente b) servidor

Figura 2. Interfaz de la aplicación

Esta aplicación multimedia, inicialmente está soportada por un middleware que proporciona las funcionalidades de conexión y transmisión/recepción de fotogramas de vídeo por la red, así como las conexiones al servidor, aceptación de clientes o empaquetado y desempaquetado de los fotogramas de vídeo. Estas funcionalidades son suministradas tanto en su versión de servidor como de cliente.

Este middleware está construido a partir de dos protocolos que implementan la transmisión de vídeo, el *Real Time Streaming Protocol, RTSP* (RFC 2326) y el *Real Time Protocol, RTP* (RFC 1889). La interacción entre el cliente y el servidor es controlada por medio del protocolo RTSP, siendo este el encargado de controlar el flujo de la información. El servidor RTSP hace un seguimiento del estado del cliente en cada sesión RTSP. Antes de que el vídeo sea enviado desde el servidor, se segmenta y se encapsula en un paquete RTP. Cuando el cliente recibe el paquete RTP, empieza su manipulación para desempaquetarlo y mostrar la información recibida.

Toda la gestión de control de flujo y transmisión de vídeo se implementa única y exclusivamente en el middleware, pero en dicho middleware no se controla el medio utilizado para la conexión de red. En el caso que la aplicación se ejecutara sobre una red inalámbrica, no se podría adaptar su comportamiento si el canal de transmisión está congestionado o si el dispositivo está perdiendo cobertura, lo que repercutiría directamente en la QoS.

Para resolver este problema, se refactorizará nuestra aplicación con el fin realizar una mejor reutilización de los elementos, aplicando la metodología AOP tal como indicamos en el apartado 3.1. Este proceso nos lleva a definir inicialmente dos aspectos:

- El DeliverySystemAspect, que es el encargado de implementar todo el sistema de reparto a través de la red.
- El RemoteVideoAspect, que enlaza a la aplicación un sistema de reparto y coordina su funcionamiento con las aplicaciones cliente y servidor.

4.1 DeliverySystemAspect

Entendemos por sistema de reparto, aquel sistema capaz de controlar los siguientes aspectos: a quién se le envía o de quién se recibe la información, qué es lo que se le envía y cuándo se le debe enviar, el cómo debe enviar o recibir la información estableciendo por dónde debe ser enviada o recibida. Además, en sí mismo un sistema de reparto no es más que un componente middleware que puede ser utilizado para diferentes usos y aplicaciones. Es por tanto independiente del resto de elementos que compongan la aplicación.

Nuestro sistema de reparto (figura 3), llamado DeliverySystemAspect está compuesto por dos elementos que desarrollan las funcionalidades básicas de cualquier sistema de reparto. El primero de ellos gestiona el control de flujo del envío y la recepción, por medio de los objetos llamados RTSPServer y RTSPClient respectivamente. Estos dos objetos son capaces de comunicarse entre sí a través de la red y utilizan RTSP como protocolo de control incorporando la máquina de estados como se especifica en la RFC 2326. El segundo elemento, encargado del envío y recepción de la información, está desarrollado por medio de los objetos RTPServer y RTPClient. Estos dos objetos, están basados en el protocolo RTP que especifica la estructura del paquete que debe usarse para poder enviar porciones de vídeo y audio por la red. RTPServer, aparte de empaquetar la información, también gestiona por dónde se va a enviar, en nuestro caso por medio de un socket de tipo UDP.

Estos dos elementos, los sistemas RTSP y RTP, son completamente independientes, pero por medio del tejedor introducimos las funciones de comunicación entre ellos, sin que ninguno de los dos anteriores sean recodificados. Esto es, el DeliverySystemAspect contiene la definición del aspecto que implanta la comunicación entre los objetos de RTSP y RTP para que trabajen conjunta, sincronizada y coordinadamente, creando así un sistema de reparto completo. Entre las tareas que debe planificar este aspecto, está el analizar el estado interno del RTSP y generar las acciones adecuadas sobre el objeto RTP.

4.2 RemoteVideoAspect

Con este aspecto se pretende aislar las comunicaciones entre las aplicaciones que necesitan de una funcionalidad de reproducción de vídeo desde fuentes remotas localizadas a través de la red, y los componentes *middleware* que implantan los mecanismos necesarios para realizar el acceso remoto. Este aspecto dispondrá por tanto de una versión cliente y otra servidor.

Los elementos que lo componen son (figura 4): a) una aplicación que dispone de la interfaz de usuario, b) un objeto capaz de suscribir los servicios de envío, recepción y control del flujo de vídeo y c) un objeto diseñado para el envío / recepción de información multimedia a través de la red.

Como puede apreciarse en la figura 4, los canales de comunicación incluídos en este aspecto son los que modelan la interconexión de la aplicación y los elementos del middleware, RTSP y RTP, que son los siguientes: 1) el establecido entre la aplicación con el elemento de suscripción que deberá fijar qué y de quién se desea recibir/prestar el servicio y 2) entre la aplicación con el elemento de envío/recepción de la información multimedia creando una pasarela entre ambos de modo que cuando se recibe una trama de vídeo desde la red, ésta es enviada al punto dentro de la aplicación que desemboca en su presentación. Es importante destacar que las comunicaciones necesarias entre los elementos a) y b) del párrafo anterior que definen y gestionan el sistema de reparto no es tarea de este aspecto, sino exclusivamente del middleware, en nuestro caso son los tejidos con el DelieverySystemAspect.

Como característica más importante de este planteamiento es que se han identificado y aislado en aspectos independientes tanto lo que es necesario desde el punto de vista de la aplicación como desde el punto de vista del middleware.

5 Aspectos para redes inalámbricas

Como se describió en el apartado 2 podemos deducir que la utilización de redes y dispositivos inalámbricos junto a las aplicaciones que se ejecutan sobre dichos dispositivos, a día de hoy, no interaccionan conjuntamente para tener un rendimiento óptimo. Los desarrollos para mejorar estos problemas pueden ser de diversa naturaleza, como la gestión del canal, control de la cobertura, optimización de recursos o control de la batería del dispositivo móvil entre otros. Nosotros hemos

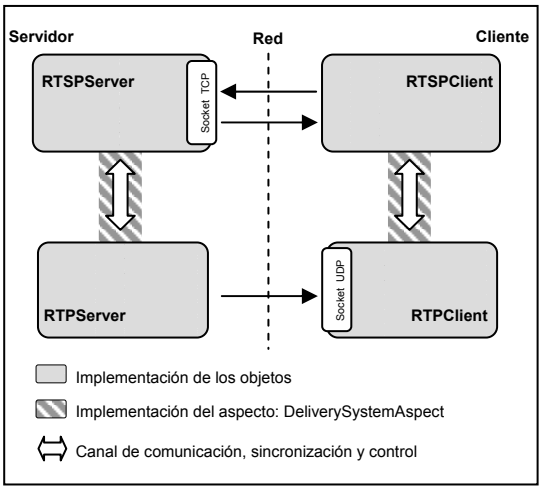


Figura 3. Esquema del DeliverySystemAspect

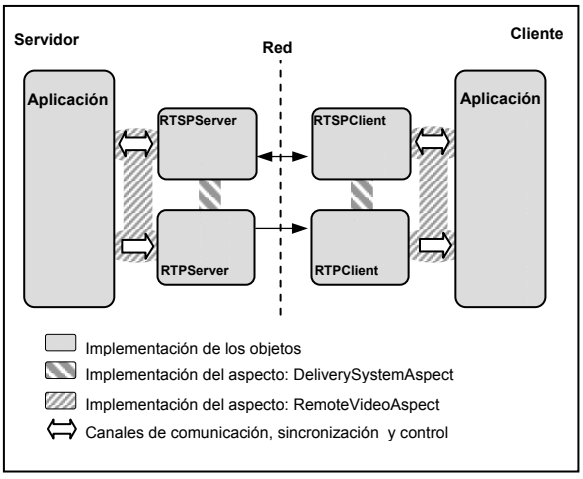


Figura 4. Esquema de la aplicación cliente y servidor tejida con el DeliverySystemAspect y RemoteVideoAspect

introducido a la aplicación multimedia dos nuevos elementos, gestión del canal y control de la cobertura, que hacen posible el tratamiento de los problemas existentes en las redes y dispositivos inalámbricos.

5.1 AvailableNetAspect

Este aspecto, cuya funcionalidad principal será la de gestionar la disponibilidad del canal, buscará información de cómo se encuentra el canal en todo momento y actuará sobre la aplicación adaptando su ejecución al estado actualizado de la red. Para realizar este procedimiento se debe tejer el aspecto con los elementos de la aplicación y los elementos de gestión del canal.

Es necesario porque la posibilidad de realizar conexiones a través de una red inalámbrica no está asegurada aún teniendo el dispositivo de red operativo. En redes cableadas las probabilidades de caídas de red son muy bajas por no decir prácticamente inexistente. Por tanto, debemos monitorizar si es posible realizar transacciones por la red, esto es, la red no está congestionada o existe cobertura suficiente. En caso de detectar que no es posible realizar transacciones por la red debemos entonces filtrar las aplicaciones que intenten hacer uso de un recurso inexistente, evitando así reintentos de conexión que a priori se sabe que van a ser infructuosos o abortar las aplicaciones que están esperando recibir datos en la red, pues estos nunca llegarán. Es más, incluso un dispositivo que tenga cobertura y posibilidad de acceso al canal, si su situación es de congestión constante, la aplicación podría cancelarse para no introducir más congestión a la red o porque no podrán alcanzarse unos parámetros de calidad de servicio satisfactorio.

En la figura 5 podemos observar los elementos que componen un aspecto de estas características:

- Objetos que nos dan la información del estado de la red: elementos agrupados en el lado derecho de la figura 5.

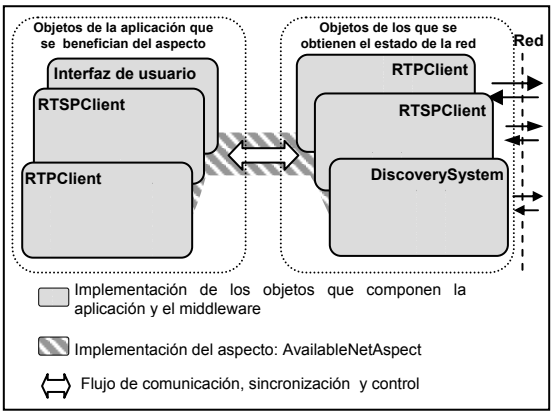


Figura 5. Esquema particular del AvailableNetAspect

- Objetos de la aplicación sobre los que puede modificarse su comportamiento en base al conocimiento proporcionado por este aspecto: elementos agrupados en el lado izquierdo de la figura 5.

- Los canales de comunicación que hacen posible tomar decisiones en base a los datos obtenidos (código del aspecto).

Existe una multitud de elementos que pueden ser insertados, esto es, tejidos junto al aspecto para que nos den información adicional del estado de la red, un ejemplo son los siguientes: Place Lab [13] (suministra clases en Java que realizan consultas al driver para obtener el estado de la red), aplicaciones tipo ping o EstarVivo (realiza operaciones de eco contra un servidor conocido que se supone disponible) o cualquier otro elemento que esté haciendo uso de la red si detectamos actividad en él. Lo importante es no introducir nuevos elementos sino ser capaz de utilizar los ya existentes para obtener esta información.

En nuestro caso estamos interesados en justificar la creación del aspecto con introducción mínima de elementos nuevos, o sea, no añadiremos ningún otro objeto que ya no está presente en la propia aplicación. Para ello nos basamos en escuchar el tráfico entre los elementos pares del tipo RTP y RTSP a través de la red y detectar cuando el canal no permite la transmisión. Por ejemplo, si estamos en el lado del cliente y se detecta que el RTP ha dejado de recibir paquetes y desconocemos si es porque no existe información multimedia o es el canal el que ya no está disponible, mediante el objeto RTSP podría detectarse si el estado de la sesión está en PLAY, en cuyo caso si solicitamos el mismo estado lo que estaremos haciendo es utilizarlo como una función ping puesto que su estado no se verá alterado y por el contrario hemos comprobado si el servidor es alcanzable.

Otro caso es si no hay red durante un intervalo de tiempo prefijado a partir del cual se puede cerrar las sesiones abiertas de RTSPServer. Las acciones a realizar sobre el sistema de reparto pueden ser de distintas características, desde no realizar ninguna acción, hasta poder parar su ejecución. Un ejemplo de lo que se puede hacer sobre el servidor del sistema de reparto sería la de poder pausar o cerrar el servicio que se está dando a un cliente ya que se observa que la red puede tener algún tipo de congestión, pudiéndose arrancar otra vez dicho servicio cuando se observe que la congestión en el canal ha disminuido lo suficiente como para poder dar un servicio óptimo.

El aspecto ahora definido entre estos dos elementos concretos, RTP y RTSP, implementa ahora una

funcionalidad que es ortogonal frente al sistema de reparto. En nuestra aplicación las acciones que se llevan a cabo cuando el aspecto recoge la información suministrada por el elemento de gestión del canal, se realizarán principalmente sobre el sistema de reparto, tanto en lo que respecta a la parte del servidor como a la del cliente.

5.2 CoverageControlAspect

Este aspecto observará la cobertura que tiene un dispositivo por medio de varios métodos como puede ser interrogando al driver del dispositivo inalámbrico, haciendo preguntas al protocolo SNMP, utilizando un sistema propietario para realizar dicha función, analizando el sistema de envío/recepción que se implementa en el propio sistema de reparto o cualquier otro que se desee que evalúe de manera efectiva el throughput de la red. Y como mencionamos en el apartado anterior, también se tendrá la aplicación multimedia con su sistema de reparto (tejido). Se debe desarrollar un aspecto que interaccione tanto con el elemento de control de cobertura como con el sistema de reparto y la aplicación. Y que dicho aspecto tome las acciones necesarias en base a la información obtenida.

En este caso las acciones a tomar son también de diversa índole, por ejemplo, si estamos hablando del sistema de reparto por parte del cliente y si el aspecto observa que la cobertura va disminuyendo, se avisa al sistema de reparto que dentro de un periodo corto de tiempo puede quedarse sin cobertura, indicándole que realice un paro en su ejecución y avisar al usuario de la aplicación por medio de mensajes que se está saliendo de cobertura y que debe moverse hacia lugares donde la potencia de la señal sea mayor. Además, en nuestro caso donde los estados del sistema de reparto están especificados por el protocolo RTSP, tanto el cliente como el servidor, podrán insertar pseudo estados que son forzados dentro de la máquina de estados del propio sistema de reparto. Estos estados "forzados" no son más que estados de PAUSE o STOP, que el aspecto gestiona automáticamente en función del estado de la cobertura siempre y cuando se estuviese en un estado de PLAY. Todos estos estados adicionales que no se corresponden con el estándar serán insertados por medio del tejedor sin que tengamos que recodificar los objetos RTSP del sistema de reparto.

6 Trabajos relacionados

No cabe duda que la aparición de diversos dispositivos móviles sobre los que se desea implantar aplicaciones es lo suficientemente atractiva para dotarlos de nuevas funcionalidades cooperativas, incluyendo las de tiempo real [21] [22]. Sin embargo, las características heterogéneas de estos dispositivos dificultan el desarrollo de aplicaciones debido a la complejidad de desarrollo que esto conlleva.

La solución a este tipo de problemas es el diseño del software con una arquitectura multinivel. Cada uno de estos niveles son soportados por uno o varios componentes middleware. Además puede afrontarse el desarrollo del middleware mediante la utilización de la AOP obteniendo beneficios frente a la metodología tradicional, como se demuestra en [9] y [23]. Por tanto, dada las ventajas de la orientación a aspectos en el desarrollo del middleware, esta se plantea como solución a entornos cambiantes [24], al permitir una mayor reutilización de código y adaptabilidad a nuevas prestaciones. Con esto se consigue reducir el tiempo y el esfuerzo de desarrollo necesario para las distintas adaptaciones.

Actualmente los antecedentes en la creación de aplicaciones cooperativas con el Diseño Orientado a Aspectos se limitan al estudio de éstos a nivel de componentes [8]. Por otro lado los desarrollos para dispositivos móviles son ínfimos [7]. Además, tenemos que el desarrollo orientado a aspectos es una tecnología novedosa, y desconocida por gran parte de los desarrolladores. Y por último, no existe una metodología que permita automatizar la creación y reutilización de aspectos, integradas en herramientas que utilicen esta tecnología para el desarrollo de aplicaciones robustas.

En [25] nosotros hemos aplicado la misma técnica de desarrollo para una aplicación que es un visor multimedia que reproduce archivos del dispositivo local. A esta aplicación se le ha tejido una serie de módulos de descubrimiento, donde el fin es descubrir la misma aplicación ejecutándose en distintas máquinas y poder cooperar remotamente en la visualización de los archivos.

7 Conclusiones

En este artículo se ha presentado el empleo de una nueva técnica para poder desarrollar aplicaciones multimedia sobre redes inalámbricas. El aprovechamiento de las ventajas que ofrece la AOP permite la reutilización de módulos ya desarrollados, por lo que se llega a un mejor nivel de modularidad y estabilidad en las aplicaciones finales. Otra característica importante de esta nueva técnica es la rapidez en desarrollar aplicaciones complejas sin introducir código redundante, y sobre todo la utilización de una metodología concreta y sistemática. Se ha desarrollado una aplicación cliente/servidor de Video Streaming basado en aspectos y la reutilización de módulos anteriormente desarrollados y probados. Tanto los módulos con los protocolos de transmisión y control de vídeo streaming como los del control de la disponibilidad de la red han sido integrados en la aplicación. Después del desarrollo de esta aplicación se ha visto que los resultados obtenidos con dicha técnica son satisfactorios, pudiendo desarrollar aplicaciones con más rapidez, más modularidad y más claridad. Al introducir módulos de control de disponibilidad de red se ha llegado a la conclusión de poder dar a los

dispositivos inalámbricos una mayor QoS, sin la necesidad de insertar nuevos módulos para el control del acceso al medio. Actualmente se está probando el nuevo módulo para el control de la cobertura del dispositivo y la definición de qué pseudo estados son los más adecuados para proporcionar mayor QoS y un mejor aprovechamiento del canal inalámbrico.

Agradecimientos

Este trabajo está subvencionado por la Consejería de Educación, Cultura y Deportes del Gobierno de Canarias (PI:164/2004).

Referencias

[1] A. Ganz, Z. Ganz, K. Wongthavarawat, *Multimedia Wireless Networks, Tecnologies, Standards and QoS*, Prentice Hall PTR, ISBN: 0-13-046099-0, 2004.

[2] A. E. Dashti, S. H. Kim, C. Shahabi, R. Simmerman, *Streaming Media Server Design*, Prentice Hall, ISBN: 0-13-067038-3, 2003.

[3] R. Gordon, S. Talley, *Essential JMF: Java Media Framework*, Prentice Hall, ISBN: 0-13-080104-6, 1998.

[4] G. Limiñana, *Desarrollo de una Herramienta Multimedia en Java sobre IP para Intercomunicación de un Grupo de Usuarios*, Departamento de Ingeniería Telemática, Proyecto Final de Carrera, Tutor: A. Suárez, EUITT- ULPGC, 2001.

[5] L. Fuentes, J. M. Troya, *A Java Framework for Web-Based Multimedia and Collaborative Applications*, IEEE Internet Computing, Vol. 3, pp. 55-64, 1999.

[6] G. Kiczales et al., *Aspect-Oriented Programming*, ECOOP'97, Jyvaskyla, Finland, Springer-Verlag 1241, junio 1997.

[7] M. Rinard, A. Salcianu, S. Bugrara, *A Classification System and Analysis for Aspect-Oriented Programs*, SIGSOFT, pp. 147-158, 2004.

[8] M. Pinto, L. Fuentes, J. M. Troya, M. Fayad, *Towards an Aspect-Oriented Framework in the Design of Colaborative Virtual Enviroment*, FTDCS'01, 2001.

[9] C. Zhang, J .Hans-Arno, *Refactoring Middleware with Aspects*, IEEE Transactions on Parallel and Distributed Systems, Vol.12, nº 11, noviembre 2003.

[10] R. M. Pratap, F. Hunleth, R. K. Cytron, *Building Fully Customisable Middleware Using an Aspect-Oriented Approach*, Proceedings of IEE, Vol.151, nº 4, pp. 199-218 , agosto 2004.

[11] Netstumbler: http://www.netstumbler.com/

[12] NDIS Developer's Reference: http://www.ndis.com/

[13] Place Lab Software: http://placelab.org/toolkit/

[14] G. Huston, *TCP in a Wireless World,* IEEE Internet Computing, Vol. 5(2), pp. 82–84, marzo/abril 2001.

[15] IEEE Xplore: Standards: http://ieeexplore.ieee.org/xpl/standards.jsp?findtitle=802.11&letter=%25802.11%25&type=2&srch=

[16] G. Tonev, V. Sunderam, R. Loader, J. Pascoe, *Location and Network Issues in Local Area Wireless Networks,* International Conference on Architecture of Computing Systems: Trends in Network and Pervasive Computing, Karlsruhe (Alemania), abril 2002.

[17] E. Macías, A. Suárez, V. Sunderam, *An Approach Toward MPI Applications in Wireless Networks,* DAPSYS, Budapest, The Kluwer International Series in Engineering and Computer Science, Vol. 777, ISBN: 0-387-23094-7, pp. 55-62, septiembre 2004.

[18] E. Macías, A. Suárez, V. Sunderam, *Efficient Monitoring to Detect Wireless Channel Failures for MPI Programs,* 12th Euromicro Workshop on Parallel Distributed and Network-Based Processing, La Coruña, pp. 374-381, febrero 2004.

[19] R. Laddad, *AspectJ in Action. Practical Aspect- Oriented Programming*, Manning Publications Co., ISBN:1-930110-93-6, 2003.

[20] G. Kiczales et al., *Getting Started with AspectJ*, Communications of the ACM, pp. 59-65, octubre 2001.

[21] A. Gal, O. Spinczyk, W. Schroder Preikschat, *On Aspect-Orientation in Distributed Real-time Dependable Systems*, 7th International Workshop on Object-Oriented Real-Time Dependable Systems, pp. 261-267, 2002.

[22] G. Singh, B. Maddula, Q. Zeng, *Enhancing Real-time Event Service for Synchronization in Object Oriented Distribuited Systems*, 7th International Workshop on Object-Oriented Real-Time Dependable Systems, pp. 233-240, 2002.

[23] C. Zhang, J. Hans-Arno, *Quantifying Aspects in Middleware Platforms*, AOSD 2003, ISBN:1-58113-660-9, pp. 130-139, Boston, 2003.

[24] V. Subramoniam, C. Gill, *A Generative Programming Framework for Adaptive Middleware*, 37th Annual Hawaii International Conference on System Sciences (HICSS'04) ISBN:0-7695-2056-1, Track 9, Vol. 9, enero 2004.

[25] S. Galván, M. Quintana, E. Mª Macías, A. Suárez, *Novel Ideas to Design Multimedia Cooperative Applications on Wireless Networks Using Aspect Oriented Programming*, Aceptado para su publicación, International Conference on Wireless Networks, Las Vegas (EEUU), junio 2005.

Educación a la carta para IDTV

Marta Rey López, Rebeca P. Díaz Redondo, Ana Fernández Vilas
Departamento de Ingeniería Telemática. Universidad de Vigo
ETSI de Telecomunicación. C/ Maxwell S/N. Campus Universitario.
36310 - Vigo (Pontevedra)
E-mail: {mrey,rebeca,avilas}@det.uvigo.es

Abstract

The great amount of contents that can be broadcast over Interactive Digital TV (IDTV), in addition to the peculiarities of the t-learning student, create the necessity of systems to personalize the learning experience in accordance with the student's preferences and knowledge, these systems are called Intelligent Tutoring Systems (ITS) in the e-learning jargon.

In this paper, we present the basis of t-*MAESTRO, an ITS for t-learning which creates personalized courses for a less formal education. We have identified two different approaches to the learning experience:* edutainment *and* entercation*, the first one closer to formal education than the other one, more related with entertainment. User modelling is the central part of the paper, since it is the principal aspect for a good performance of a personalization system.*

1. Introducción

El término *t-learning* se utiliza con el significado de aprendizaje interactivo a través de un televisor [13]. No es simplemente una adaptación de *e-learning* para IDTV, pues cuenta con sus propias características distintivas, relacionadas, en gran parte, con las restricciones impuestas por el televisor y el *set-top box*, como la baja resolución de la pantalla, el hecho de que la interacción haya de llevarse a cabo a través de un mando a distancia o las bajas prestaciones de un *set-top box* —mucho más limitadas que las de un ordenador.

Igualmente importantes resultan otras particularidades que no tienen que ver con aspectos tecnológicos sino sociales: la predisposición que un alumno de *t-learning* presenta hacia la educación es completamente diferente a la de aquél de *e-learning*. Mientras que el segundo tiene una actitud activa hacia el aprendizaje, ya que ha sido él mismo quien ha decidido tomar parte en la experiencia educativa; el primero es normalmente más pasivo y habrá de ser atraído hacia ella a partir de actividades de entretenimiento que puedan resultarle interesantes.

Dicha pasividad —herencia de los más de 50 años de televisión analógica, donde la interacción entre telespectador y televisor se reduce a un mando a distancia—, unida a la concepción de la televisión como un medio dedicado al entretenimiento, hacen necesario entender que el telespectador no buscará, por lo general, una educación formal sino que se verá involucrado en una experiencia educativa a partir del entretenimiento.

Sin embargo, no todo son obstáculos para la educación a través de IDTV, ya que existen diferentes razones que la convierten en un medio adecuado para ello [13]: por una parte, alrededor de un 98 % de los hogares europeos cuentan con al menos un televisor, por otra, la gente tiende a confiar en todo aquello que ve en televisión —condición indispensable para la educación.

Además de los factores ya mencionados, la IDTV permite la difusión de un mayor número de canales que la televisión analógica tradicional, posibilitando una mayor variedad de contenidos audiovisuales y educativos. Este hecho, si bien es claramente una de sus mayores ventajas, trae consigo la necesidad de herramientas especializadas que permitan localizar aquellos contenidos que realmente interesan al telespectador, puesto que la ingente cantidad de información a la que un usuario tiene acceso puede llegar a abrumarlo [12].

Estas herramientas habrán de satisfacer no sólo la necesidad de selección de contenidos educativos, sino también la de personalización de los mismos, haciendo que el aprendizaje resulte más atractivo y efectivo. Más atractivo porque se adaptará a las preferencias del alumno y más efectivo pues tendrá en cuenta la experiencia formativa del estudiante y las metas que aspira a alcanzar. Los sistemas que llevan a cabo esta tarea se conocen como *Intelligent Tutoring Systems* (ITS).

En este artículo, presentamos un ITS para el campo de *t-learning*, capaz de adecuar los contenidos educativos existentes a cada usuario, mejorándolos con recursos que han sido difundidos como contenido audiovisual. Por semejanza con el término *t-learning*, hemos llamado a este sistema *t*-MAESTRO (**M**ultimedia **A**daptive **E**ducational **S**ys**T**em based on **R**eassembling TV **O**bjects) y utilizaremos nuestra experiencia previa en el recomendador de contenidos audiovisuales AVATAR [2] para su desarrollo.

El texto se organizará de la siguiente manera: en los apartados 2, 3 y 4 presentamos las aproximaciones posibles a la experiencia educativa a través de *t*-MAESTRO, su modo de funcionamiento

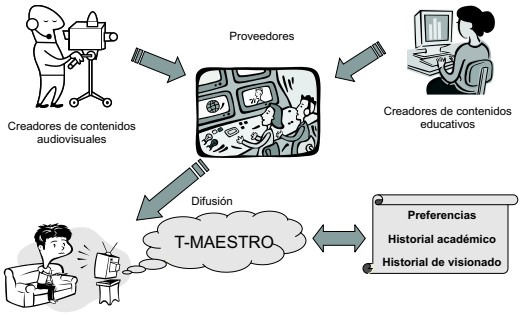


Figura 1: Escenario de *t-learning*

y los principales elementos que lo componen, dedicando el apartado 5 al modelo de usuario, elemento central de todo sistema de personalización. Por último, el apartado 6 propone una arquitectura acorde a los elementos presentados y en el apartado 7 discutiremos la relevancia del trabajo y presentaremos futuras líneas de investigación.

2. Escenario

Antes de describir las características propias de *t*-MAESTRO, se hace necesaria la identificación del escenario sobre el que va a trabajar. En este apartado, describiremos, entre otros, las características del usuario y el tipo de contenidos que serán difundidos.

En un futuro estable de la IDTV, los proveedores difundirán contenidos audiovisuales creados por terceros. Adicionalmente, tendremos creadores de contenidos educativos, que serán difundidos junto a los audiovisuales. De este modo, ambos tipos de contenido llegarán al *set-top box* del cliente, donde reside *t*-MAESTRO (Fig. 1), quien filtrará y combinará los contenidos recibidos —tanto audiovisuales como educativos— con el objetivo de crear cursos personalizados de acuerdo con información almacenada acerca del usuario (sus preferencias, historial, experiencia formativa, etc.).

Como acabamos de comentar, la finalidad última de *t*-MAESTRO consiste en combinar y organizar un subconjunto de los contenidos que recibe y que identifica como interesantes para el curso a crear. Las situaciones que podríamos encontrar son variadas y, tal como sucede en la televisión analógica tradicional, el cliente podría recibir en su *set-top box* contenidos audiovisuales como los ya conocidos, es decir, programas de televisión. Acompañando a dichos programas podría encontrarse contenido educativo relacionado con el tema tratado en los mismos, que permita al usuario profundizar en la materia. Por otra parte, algunos canales temáticos podrían difundir contenido educativo relacionado con sus objetivos de forma paralela al audiovisual, e incluso podrían existir canales dedicados única y exclusivamente a fines educativos.

En este escenario tecnológico concebimos dos predisposiciones a la educación a través de la televisión. Partiendo de que el estudiante de *t-learning* tiene muy probablemente una actitud pasiva hacia la educación, ha surgido un nuevo concepto, *edutainment*, abogando por educación (*education*) y entretenimiento (*entertainment*) y refiriéndose a la aproximación adecuada para el desarrollo de modelos para un proceso de aprendizaje menos formal [4]. Desde nuestro punto de vista, las experiencias de *edutainment* se basan en el seguimiento de un curso de naturaleza educativa formal que ha sido mejorado para que resulte más efectivo y entretenido añadiéndole contenido audiovisual relacionado.

Con este modelo, estamos suponiendo que el telespectador tiene una cierta predisposición al aprendizaje, sin embargo la creación de cursos para usuarios más pasivos que no tengan ese interés por el aprendizaje es también posible. Consiste en ofrecer al telespectador contenidos educativos relacionados con el programa que está viendo y que le permitirán profundizar en el tema sobre el que éste versa. Por ello, si decide participar en la experiencia educativa que se le ofrece, lo hará sin tener la más mínima intención previa. Hemos llamado *entercation*[1] a este modo de aprendizaje, puesto que es más próximo al entretenimiento que a la educación formal.

3. Descripción general

En el apartado anterior presentamos el escenario en que ubicaremos *t*-MAESTRO, en éste, describimos el funcionamiento del mismo de forma general. A continuación —con ayuda de la Fig. 2— exponemos los diferentes pasos que dan lugar a la creación de cursos personalizados.

1. El flujo de transporte se filtra para obtener los programas de televisión —para el procesamiento de sus metadatos por AVATAR— y el contenido educativo —para el Recomendador educativo. Este filtrado se lleva a cabo de acuerdo con el etiquetado de los contenidos, de modo que, si los contenidos se etiquetan para ser usados como elementos educativos[2] además de como programas de televisión, serán enviados hacia ambos.

2. El siguiente paso consiste en la selección de contenidos: AVATAR selecciona los programas de televisión que considera interesantes para el usuario según su perfil como telespectador (es importante mencionar que seleccio-

[1] Si bien este término ya había sido utilizado anteriormente de forma esporádica, no llevaba consigo el significado de contraposición a *edutainment* que nosotros aportamos.

[2] En terminología propia de SCORM, el estándar de *e-learning* que seguiremos, la mínima unidad de contenido educativo que puede mostrarse al usuario es conocida como SCO (apartado 4.2).

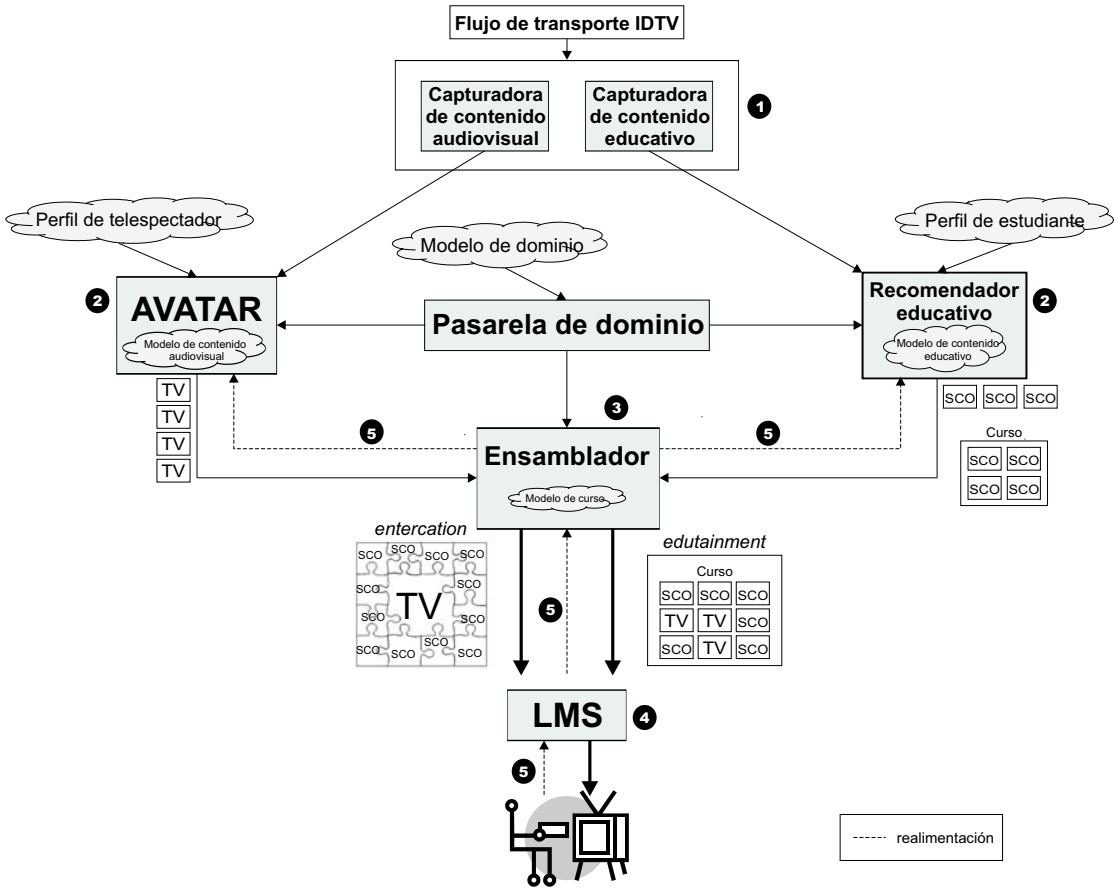


Figura 2: Funcionamiento de *t*-MAESTRO

na los mismos programas que cuando funciona de forma independiente a *t*-MAESTRO), mientras que el Recomendador educativo selecciona los contenidos educativos que encajan en el perfil de estudiante, teniendo en cuenta su nivel, experiencia educativa previa, etc. Ambos hacen uso de la Pasarela de dominio, subsistema basado en razonamiento semántico que permite relacionar conceptos más allá de la simple coincidencia sintáctica.

3. El Ensamblador crea el curso personalizado, haciendo uso de un modelo de curso, que contendrá reglas acerca de cómo combinar los objetos seleccionados para obtener un curso estructurado en lugar de un conjunto de objetos aislados.

4. El asistente de ejecución —LMS (*Learning Management System*) en terminología de estándares de *e-learning*— recibe el curso creado y lo presenta al usuario.

5. La realimentación que aporta el usuario al sistema hace posible la actualización continua de los perfiles de estudiante y telespectador.

Como resultado, el usuario obtendrá una experiencia educativa de *edutainment* —es decir, cursos más formales enriquecidos con contenido audiovisual— o *entercation* —educación menos formal basada en complementar el contenido audiovisual recomendado al telespectador con elementos educativos relacionados.

Por ejemplo, supongamos que el estudiante para el que vamos a crear los cursos es un hombre de negocios con las siguientes características: le gustaría mejorar su inglés, tiene almacenados en su historial de visionado numerosos partidos de fútbol y está interesado en la salud relacionada con el deporte. En el caso de *edutainment*, *t*-MAESTRO podría ofrecerle un curso de inglés a medida combinando los siguientes: "Inglés en los negocios", "Particularidades del inglés americano" y "Pronunciación del inglés", enriqueciéndolo con un partido de fútbol retransmitido en este idioma para que mejore su comprensión oral. *t*-MAESTRO también podría ofrecerle una experiencia de *entercation*, por ejemplo, si AVATAR ha seleccionado un partido de fútbol para el usuario, *t*-MAESTRO podrá localizar recursos educativos relacionados con ese partido, como un documental acerca de los peligros de la vida sedentaria, una aplicación que muestre cómo mejorar tu salud haciendo deporte y un vídeo interactivo cuya temática sea la prevención de lesiones.

4. Principales Elementos

En el sistema que estamos describiendo, existirán tres elementos fundamentales que habremos de modelar: el dominio —que permitirá establecer relaciones entre conceptos—, el contenido —tanto audiovisual como educativo— y el usuario —elemento central a la hora de personalizar, que comprende las facetas de telespectador y estudiante. Describiremos los dos primeros en este mismo apartado y dedicaremos el apartado 5 al modelo de usuario.

Además de describir dichos elementos, destacaremos la importancia de la normalización de los mismos, tomando como punto de partida estándares existentes predominantes en el campo correspondiente. La normalización permitirá la reutilización de elementos tanto educativos como audiovisuales para diferentes cursos creados por *t*-MAESTRO, así como que estos cursos sean independientes del hardware y software que el usuario tenga en su *set-top box*.

4.1. Modelado de dominio

Como ya se mencionó en el apartado 2, ha de existir un subsistema que establezca relaciones entre conceptos relativos al mismo dominio. Para que esto sea posible, los conceptos de un dominio estarán almacenados en una ontología, que constituye una especificación explícita de una conceptualización compartida. En [14], se propone la utilización de ontologías en el contexto educativo para permitir la reutilizabilidad e interoperabilidad.

El modelo de dominio estará constituido por un conjunto de ontologías —cubriendo los temas de interés del usuario— distribuidas e interrelacionadas, en las que se buscarán los vínculos existentes entre diferentes conceptos. Esto permitirá establecer correspondencias entre el perfil de usuario y el contenido más allá de la simple coincidencia sintáctica.

Proponemos la utilización del lenguaje OWL (*Web Ontology Language*) [17] para la creación de las citadas ontologías, ya que ha sido diseñado específicamente para que sea compatible con la web semántica y está siendo ampliamente adoptado.

4.2. Modelado de contenido

Hemos de diferenciar dos clases de contenido: a un lado encontramos el contenido puramente educativo, creado para ese propósito, que consistirá en cursos estructurados acerca de un determinado tema, así como de piezas aisladas de contenido educativo; al otro lado, el contenido audiovisual tradicional —cuyo propósito principal es el entretenimiento.

En lo referente al **contenido audiovisual**, éste se difundirá de acuerdo con la norma DVB-MHP [5], que utiliza flujos de transporte MPEG-2 para empaquetar y multiplexar los contenidos. En concreto, las aplicaciones se montan en una estructura conocida como Carrusel de Objetos [10], que se transmite repetida y periódicamente, ofreciendo un sistema local de ficheros.

Para la descripción de los programas de televisión, supondremos que sus creadores siguen la especificación *TV-Anytime Metadata* [15], que permite etiquetar dichos programas —o segmentos de ellos— con metadatos cuya parte más visible es aquella que encontramos en las EPGs (*Electronic Program Guides*) describiendo el contenido.

En cuanto al **contenido educativo**, después de estudiar varios estándares en el campo de *e-learning*, hemos decidido que el más adecuado es ADL SCORM (*Advanced Distributed Learning - Sharable Content Object Reference Model*) [1], puesto que toma como base importantes estándares de *e-learning* y crea guías para su uso conjunto.

Para la descripción de cada uno de los elementos que componen un curso, SCORM utiliza los metadatos definidos por la especificación IEEE LOM (*Learning Object Metadata*) [8]. Estos cursos son empaquetados en un archivo PIF (*Package Interchange File*), que, según nuestra propuesta, ha de ser enviado a través del carrusel de objetos mencionado con anterioridad. Para mostrar los cursos al usuario una vez recibidos dichos ficheros, SCORM define un sistema capaz de gestionar dichos contenidos: el LMS (*Learning Management System*). La unidad mínima de contenido educativo que éste podrá manejar se conoce como **SCO** (*Sharable Content Object*).

En el sistema que proponemos, los metadatos asociados a los programas de televisión —o segmentos de los mismos— podrán ser ampliados con metadatos SCORM si su creador considera que éstos podrían ser usados no sólo para el entretenimiento sino también con fines educativos. Tal como se discute en [6], es posible la transformación entre metadatos *TV-anytime* y metadatos SCORM para la creación de cursos compatibles con este último que utilicen además contenido audiovisual con fines de entretenimiento.

5. Modelado de usuario

El modelo de usuario constituye el núcleo de los sistemas personalizados, ya que supone el elemento alrededor del cual giran los algoritmos que permiten crear cursos personalizados para cada usuario concreto.

Tal como se define en [11], existen tres modos diferentes de personalizar un sistema, que nosotros particularizaremos para el caso concreto que estamos describiendo:

Personalización de la estructura Se refiere al modo en que se estructura y presenta el conjunto de objetos que conforman el contenido según los diferentes usuarios.

Adaptación del contenido Proceso de adecuar dinámicamente qué información se presenta a los diferentes usuarios de acuerdo con sus perfiles.

Adaptación de la presentación y el modo Permite a los usuarios elegir entre diferentes estilos de presentación, aunque también posibilita cambios en el tipo de medio que se va a presentar al usuario, el estilo de aprendizaje, etc. (adaptación de modo).

En nuestro caso concreto, estamos especialmente interesados en los dos últimos tipos de personalización: la personalización del contenido es el objetivo principal de *t*-MAESTRO, si bien la adaptación de modo nos resulta igualmente interesante, ya que permite que si un concepto puede enseñarse a través de diferentes medios —audio, vídeo, texto...—, se muestre al usuario sólo aquél por el que haya demostrado preferencia.

5.1. Datos almacenados

Los datos almacenados en el perfil de usuario pueden clasificarse en tres categorías bien definidas: **datos de usuario** —información personal acerca de las características del usuario—, **datos de uso** —que resultan de la interacción del usuario con *t*-MAESTRO— y **datos de entorno** —como puede ser el hardware y software que posee el usuario o el idioma que utiliza.

Mientras que los datos de usuario y entorno se obtienen de forma estática a través de preguntas directas, los datos de uso se obtienen de forma dinámica, por ello, el sistema ha de valerse de servicios que permitan actualizar el perfil de usuario a partir de los datos recogidos por medio de la interacción del usuario con el sistema.

En nuestro caso concreto, el usuario de los servicios de *t*-MAESTRO será el propio telespectador, por lo que tendremos dos modelos de usuario distintos: el modelo de telespectador —creado por AVATAR, que almacenará la información relativa a sus preferencias como televidente— y el modelo de estudiante —que será gestionado por el Recomendador educativo. Es importante mencionar que en este último no se almacenará solamente el nivel de conocimiento que el usuario tiene en un determinado dominio, sino también las preferencias de éste hacia temas concretos. Ambos aspectos serán tratados en los apartados 5.2 y 5.3, respectivamente.

Para almacenar el modelo de estudiante, nos basaremos en uno de los estándares ya existentes: IMS LIP (*Learner Information Package*) [9], inspirado inicialmente en IEEE PAPI (*Public and Private Information*) [7], provee una gran flexibilidad ya que, por una parte, la mayoría de sus elementos son opcionales y, por otra, es posible extenderlos para adecuarlos a nuestras necesidades.

LIP define once categorías, que han de ser adaptadas a las peculiaridades del alumno de *t-learning*, ya que sus características son diferentes a las de *e-learning*. A continuación describimos dichas categorías y cómo las hemos adaptado nosotros para aplicarlas a un entorno de *t-learning*:

Para la correcta identificación del usuario de *t-learning* se utilizan las categorías ***Identification*** y ***Security Key***, que permiten el acceso al sistema mediante autenticación con contraseña.

Activity permite almacenar en qué actividades de aprendizaje ha participado el alumno y en qué medida éstas han sido productivas para su bagaje académico.

Competency almacena las habilidades, conocimientos y capacidades que ha adquirido el estudiante en actividades de aprendizaje llevadas a cabo en *t*-MAESTRO o fuera de él (apartado 5.2).

Accessibility incluye las preferencias de usuario, tanto personales como tecnológicas. Ésta es una categoría esencial en *t-learning* ya que indicará sus preferencias de presentación y modo, tales como el lenguaje, el tipo de contenidos que prefiere o su canal de difusión preferido.

Interest guarda información que describe los pasatiempos y actividades de tiempo libre del estudiante, incluyendo las preferencias de estudiante (apartado 5.3). Esta categoría es clave para complementar los contenidos audiovisuales seleccionados por AVATAR con unidades pedagógicas atractivas para el usuario (objetivo de *entercation*).

Por otro lado, la categoría ***Goal*** supone la base para la obtención de un curso de *edutainment* dado que en ella se almacenan los objetivos de aprendizaje del usuario.

Las categorías ***QCL***, ***Transcript*** y ***Affiliation*** —que contienen, respectivamente, calificaciones, información académica y pertenencia a organizaciones profesionales del alumno— son aplicables sólo en caso de que el alumno sea activo y disponga de un dispositivo que permita una mayor interacción que un mando a distancia tradicional. Tienen sentido en una educación más formal, por lo que podrían ser completadas por un estudiante de *edutainment* pero no por uno de *entercation*.

Hasta el momento, hemos visto las distintas categorías de un modo aislado, sin embargo, la categoría ***Relationship*** permite expresar las posibles relaciones existentes entre ellas. En *t-learning* mostrará, entre otras, las relaciones entre las actividades en las que el estudiante ha tomado parte y la información obtenida a partir de ellas, por ejemplo, que prefiere las aplicaciones a los documentales.

En estas categorías, habremos de almacenar los datos de usuario, uso y entorno citados en 5.1. Los datos de entorno serán almacenados en las categorías, *accessibility* y *security key*, los datos de usuario se almacenarán en las categorías *identification*, *interest*, *goal*, *QCL*, *transcript* y *affiliation*, mientras que los datos de uso se guardarán en *activity*, *competency*, *accessibility* e *interest*.

5.2. Conocimientos del estudiante

Para almacenar los conocimientos del alumno acerca de los conceptos de un determinado dominio, utilizaremos un modelo multicapa [16], de modo que cada concepto perteneciente a un dominio puede alcanzar cuatro niveles diferentes de aprendizaje: **visitado** —el usuario ha llevado a cabo alguna actividad referente a ese concepto—, **aprendido** —el usuario ha demostrado mediante algún procedimiento de evaluación que ha aprendido dicho concepto—, **inferido** —el sistema ha inferido a partir de otra información que el alumno ha asimilado el concepto— y **conocido** —el estudiante ha indicado al sistema que ya conoce el concepto.

Supongamos que los conceptos relativos a un dominio en concreto (aquéllos definidos en la ontología de dicho dominio) constituyen un vector de longitud n, tendremos una matriz $4 \times n$ en la que cada columna representará el grado de conocimiento acerca de un determinado concepto. Dicha matriz —que llamaremos **Matriz de conocimiento**— representará el conocimiento de un alumno en un determinado dominio. Por ejemplo, el conocimiento de un dominio cuyos conceptos son $c_1 \ldots c_{10}$ podría ser el siguiente:

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
aprendido	NO	NO	NO	NO	SÍ	NO	NO	NO	NO	NO
inferido	NO	NO	SÍ	NO	NO	NO	NO	NO	SÍ	NO
conocido	NO	SÍ	NO	SÍ	NO	NO	NO	SÍ	NO	SÍ
visitado	SÍ	NO	NO	NO	NO	NO	SÍ	NO	NO	NO

Uno de los grandes problemas que se presentan a la hora de construir el perfil de usuario es la inicialización del mismo, la opción más sencilla para resolver este problema consiste en suponer que todo usuario inicia su aprendizaje en un dominio con conocimiento nulo, o con un conocimiento predeterminado, igual para cualquier usuario. Esta solución no es adecuada para un entorno de personalización debido a la heterogeneidad de sus usuarios, más patente si cabe en el campo de *t-learning*, por ello recurriremos a la creación de **estereotipos de conocimiento**.

Cuando un usuario comienza a aprender acerca de un determinado dominio, no sabemos su nivel de conocimiento en el mismo, sin embargo, podríamos tener una base de datos con matrices de conocimiento estereotipo, una para categoría diferente de usuario que se haya definido. Cada usuario pertenecerá a una categoría en función de la información de la que se disponga acerca de él y su matriz de conocimiento se inicializará con la matriz de conocimiento estereotipo perteneciente a dicha categoría. Es importante destacar que los conceptos del nuevo dominio que tomamos de los estereotipos, sólo podrán tener el nivel de inferidos, pues en ningún caso sabremos con seguridad el conocimiento que sobre estos conceptos tendrá el usuario, sino que lo habremos inferido a partir de la información procedente del resto de usuarios que pertenecen a esta categoría.

Pueden existir conceptos que pertenezcan a varios dominios y por ello, puede que ya tengamos información sobre algunos conceptos del nuevo dominio relativa a nuestro usuario. Si esto sucede, las columnas correspondientes a estos conceptos serán inicializadas con la información previa y no con aquella presente en la matriz estereotipo. Para identificar estos conceptos que suponen un puente entre ontologías de dominios diferentes, nuestro sistema se valdrá de la Pasarela de dominio de la que se ha hablado en el apartado 3.

5.3. Preferencias de estudiante

Hasta el momento, hemos hablado solamente del conocimiento de un usuario acerca de un dominio, sin embargo, éste no es el único factor de importancia a la hora de crear un curso personalizado, hemos de saber, además, si dicha temática le interesa o no. Para ello, el sistema cuenta con información de preferencias en dos capas.

En primer lugar, la interacción del usuario con el sistema, dará lugar a las **preferencias de estudiante**, que contendrán no sólo información estática directamente introducida por éste en el momento en que se registra en el sistema, sino también otra información dinámica que el sistema ha inferido (o simplemente, conocido) a partir de sus acciones e historial.

En segundo lugar, en un nivel superior, encontraríamos los **estereotipos de preferencias**, que consisten en información acerca de preferencias de estudiantes genéricos. El gestor de estereotipos será un agente encargado de revisar las preferencias de los distintos estudiantes que pertenezcan al sistema (por ejemplo, aquéllos que obtengan los servicios de *t-learning* de un mismo distribuidor) y crear estereotipos.

Para saber si un determinado contenido interesa al estudiante, el Recomendador educativo consulta primero las preferencias de estudiante, si encuentra alguna información sobre el tema, tanto positiva como negativa (es decir, podría indicar que el usuario está interesado en ese tema o que no le interesa en absoluto), la tiene en cuenta. En caso de no se aporte información acerca de ese tema en concreto, consultará los estereotipos de preferencias, para ver el grado de interés de estudiantes con perfiles similares y asumirá que su usuario concreto va a tener los mismos gustos.

6. Arquitectura propuesta

Partiendo de toda la información aportada acerca de *t*-MAESTRO en apartados anteriores, mostramos ahora en la Fig. 3 la arquitectura en la que encaja todo ello. En esta figura mostramos los diferentes subsistemas que componen *t*-MAESTRO, que estarán compuestos por diferentes servicios, representados por elipses. El flujo de datos entre

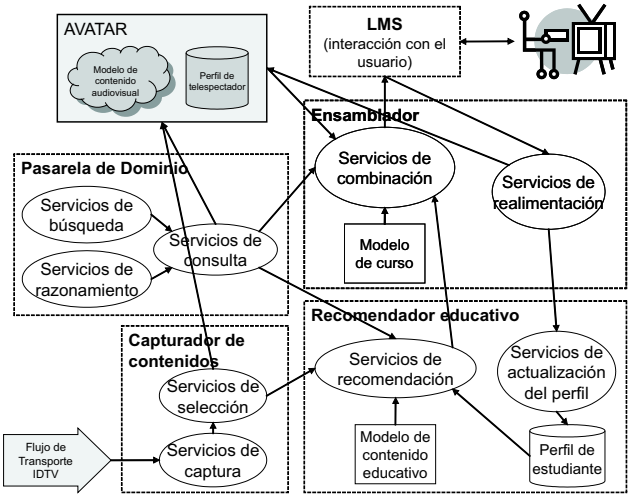


Figura 3: Arquitectura del sistema

servicios se representa en la figura por medio de flechas, que indican además el sentido de esta transferencia de datos.

Capturador de contenido. Su misión es la captura y selección del contenido. El flujo de transporte pasa a través de los **Servicios de captura**, que recuperan los contenidos y los envían a los **Servicios de selección**, quienes los distribuyen al Recomendador educativo o a AVATAR, según corresponda.

Pasarela de dominio. Este sistema es el encargado de establecer relaciones semánticas entre los términos pertenecientes a un determinado ámbito, con el objetivo de facilitar una búsqueda de contenidos conformes al perfil correspondiente. Necesita los **Servicios de consulta** para recibir y responder las consultas, quienes invocan a los **Servicios de búsqueda** para encontrar la ontología del dominio sobre el cual se realiza la consulta y a los **Servicios de razonamiento**, cuya misión consiste en establecer las relaciones pertinentes.

Recomendador educativo. Selecciona los contenidos educativos que se adaptan al perfil de estudiante. Para llevar a cabo esta tarea utiliza los **Servicios de recomendación**, quienes tienen en cuenta el perfil de estudiante, actualizado por los **Servicios de actualización del perfil** de acuerdo con los cursos y contenidos educativos seleccionados por el usuario.

AVATAR. Se encarga de seleccionar los contenidos audiovisuales que se adaptan al perfil de telespectador según la arquitectura descrita en [2].

Ensamblador. Analiza el contenido seleccionado por AVATAR y el Recomendador educativo combinándolo de forma estructurada para crear cursos. Los **Servicios de combinación** llevan a cabo la combinación de elementos, utilizando los modelos de contenido (educativo y audiovisual) y el modelo de curso, que consistirá —para el caso de *edutainment*— en aquél definido por SCORM y que aún está por definir para *entercation*. Además, los **Servicios de realimentación** reciben la información que aporta el usuario y deciden a qué subsistema —AVATAR o el Recomendador educativo— puede resultar útil dicha información.

LMS. Tal como se define en SCORM, es el encargado de gestionar los cursos para que el usuario pueda verlos e interactuar con él para permitir, por ejemplo, la actualización de sus perfiles.

7. Conclusiones y Trabajo Futuro

En este artículo, hemos presentado las bases de *t*-MAESTRO, un tutor inteligente en el campo de *t-learning*. Su característica más destacable radica en la posibilidad de crear cursos a la carta adaptados a las necesidades de cada alumno concreto, más allá de la simple selección de cursos genéricos. Otra novedad consiste en la inclusión de contenidos audiovisuales —que no han sido específicamente creados con fines educativos— en los propios cursos, como material de apoyo a los contenidos educativos. Esta última característica permite que los cursos resulten más entretenidos y atractivos para el estudiante, moviéndonos, de este modo, en los dominios de *edutainment* y *entercation* —término que hemos acuñado para referirnos a una educación poco formal basada en el entretenimiento— y no en los de la educación puramente formal. Los cursos de *edutainment* se ofrecerán al usuario como un todo en el momento de la recomendación de material audiovisual, mientras que los cursos de *entercation* consistirán en material educativo de apoyo a aquellos contenidos audiovisuales recomendados al usuario y que éste puede decidir cursar o no.

Por otra parte hemos destacado la importancia de la reutilizabilidad de contenidos y la interoperabilidad entre sistemas, dos principios que, según la experiencia acumulada en *e-learning*, resultan cruciales para la reducción de los costes de desarrollo. Estos dos principios constituyen los pilares

sobre los que se sustenta *t*-MAESTRO, ya que, de otro modo, sería impracticable conseguir la funcionalidad propuesta.

La utilización de estereotipos puede abrir las puertas a dos líneas futuras. En primer lugar podemos compartir los cursos creados para un usuario concreto con otros pertenecientes a su misma categoría. Por otra parte, es muy común que un mismo televisor se utilice a la vez por múltiples usuarios —por ejemplo, los miembros de una familia—, por lo que sería interesante crear un curso que se adaptase a las necesidades de varios usuarios, agrupando sus perfiles en uno solo de acuerdo a sus similitudes. Para llevar a cabo esta tarea, sería necesario añadir un sistema de negociación que busque elementos comunes en preferencias y conocimiento.

Por último, se hace necesaria la definición de un modelo de curso para *entercation*, puesto que al ser una nueva concepción de la educación, no son válidos los modelos ya definidos. Si bien hay diferencias conceptuales entre ambos, podremos partir de la idea de hipermedia adaptativo ampliamente utilizada en *e-learning* [3].

Agradecimientos

Este trabajo ha sido parcialmente subvencionado por el Ministerio de Educación y Ciencia, proyecto de investigación TSI 2004-03677.

Referencias

[1] ADL. Sharable Content Object Reference Model (SCORM). `http://www.adlnet.org`, 2004.

[2] Y. Blanco Fernández, J. J. Pazos Arias, A. Gil Solla, M. Ramos Cabrer, B. Barragáns Martínez, M. López Nores, J. García Duque, A. Fernández Vilas, and R. P. Díaz Redondo. AVATAR: An advanced Multi-Agent Recommender System of Personalized TV Contents by Semantic Reasoning. In S. Verlag, editor, *Web Information System Engineering*, LNCS, pages 415–421, 2004.

[3] P. Brusilovsky. Methods and techniques of adaptive hypermedia. *User Modeling and User Adapted Interaction (Special Issue on adaptive hypertext and hypermedia)*, 6(2-3):87.129, 1996.

[4] D. Buckingham and M. Scanlon. *Education, Entertainment and Learning in the Home.* Open University Press, 2002.

[5] DVB Consortium. Multimedia Home Platform Specification 1.2.1. European Standard ETSI TS 102 812 V1.2.1, 2003.

[6] M. Frantzi, N. Moumoutzis, and S. Christodoulakis. A Methodology for the Integration of SCORM with TV-Anytime for Achieving Interoperable Digital TV and e-Learning Applications. In *Proceedings of the IEEE International Conference on Advanced Learning Technologies (ICALT'04)*, 2004.

[7] IEEE Learning Technology Standards Committee (LTSC). Private and Public Information. `http://edutool.com/papi/`, 2001.

[8] IEEE Learning Technology Standards Committee (LTSC). Learning Object Metadata. IEEE Standard 1484.12.1, 2002.

[9] IMS Global Learning Consortium. Learner Information Package (LIP). `http://imsproject.org`, 2001.

[10] International Organization for Standardization ISO/IEC. Information Technology - Generic Coding of Moving Pictures and Associated Audio Information: Extensions for Digital Storage Media Command and Control. ISO/IEC 13818-6, 1998.

[11] A. Kobsa, J. Koenemann, and W. Pohl. Personalised hypermedia presentation techniques for improving online customer relationships. *The Knowledge Engineering Review*, 16(2):111–155, 2001.

[12] M. López Nores, J. J. Pazos Arias, Y. Blanco Fernández, M. Rey López, J. García Duque, B. Barragáns Martínez, A. Fernández Vilas, R. P. Díaz Redondo, A. Gil Solla, and M. Ramos Cabrer. Solutions for personalized t-learning. In *3rd European Conference on Interactive Television: User Centred ITV Systems, Programmes and Applications (EuroITV-05).*, 2005.

[13] pjb Associates. A Study into TV-based Interactive Learning to the Home. `http://www.pjb.co.uk/t-learning`, 2003.

[14] J. Qin and N. Hernández. Ontological Representation of Learning Objects: Building Interoperable Vocabulary and Structures. In ACM Press, editor, *Proceedings of WWW2004*, pages 348–349, 2004.

[15] The TV-Anytime Forum. Broadcast and Online Services: Search, select and rightful use of content on personal storage systems. European Standard ETSI TS 102 822, 2004.

[16] G. Weber and P. Brusilovsky. ELM-ART: An Adaptive Versatile System for Web-based Intruction. In *Proceedings of International Journal of Artificial Intelligence in Education*, volume 1, pages 351–384, 2001.

[17] World Wide Web Consortium (W3C). Web Ontology Language (OWL). `http://www.w3.org/2004/OWL/`, 2004.

Arquitectura de Servicios Basados en Servlets SIP[1]

Jesús Alcober Segura*, Sergio Machado Sánchez *, Antonio Oller Arcas*, Xavier Hesselbach i Serra*, Antonio Abajo Álvarez, Gines Gómez†, Jesús Rodríguez †
*Departamento de Ingeniería Telemática. Universidad Politécnica de Cataluña, . †Voztele.com
Avda. del Canal Olímpico S/N.
08860 – Castelldefels (Barcelona)
Teléfono: 93 413 72 06 Fax: 93 413 70 07
E-mail: {jalcober, smachado, aoller, xavier.hesselbach}@entel.upc.edu

***Abstract**. We present a services architecture based on SIP Servlets for the development of enhanced services in SIP networks. The SIP Servlet API is a specification of the Java community that defines a high-level API for SIP servers, enabling SIP applications to be deployed and managed based on the well known HTTP Servlet model. We also present an implementation of this specification that we have developed for the IP telephony operator VozTelecom, discussing its design and performance results. Finally, we also present a framework for the development of SIP services based on the design pattern of state machines.*

1 Introducción

Actualmente, la evolución en el entorno de desarrollo de servicios de Internet se centra en facilitar el proceso de diseño e implementación de las aplicaciones que ofrecen servicios. Ya desde los primeros tiempos de la popularización de Internet hubo múltiples propuestas en cuanto a la aplicación de estrategias que permitiesen alcanzar un sistema capaz de gestionar comunicaciones utilizando los protocolos pertinentes. HTTP (*HyperText Transfer Protocol*) es el protocolo básico para acceder a la información disponible en la Web. La evolución del entorno de aplicación de dicho protocolo supuso un aumento en la complejidad del desarrollo de las aplicaciones a medida que las posibilidades iban incrementándose.

La simplificación del desarrollo de estas aplicaciones comenzó con la aparición de los CGI (*Common Gateway Interface*), un estándar para conectar servidores de información como los servidores Web a aplicaciones externas. Así, una página Web servida por un *daemon* Web es estática, sin embargo, si previamente a la respuesta a la petición HTTP el servidor ejecuta cierta aplicación externa la respuesta obtenida por el cliente puede ser dinámica, añadiendo cierta "inteligencia" al puro servicio de acceso a documentos alojados en la Web. La idea original de los CGI ha ido evolucionando en diversas tecnologías hijas como PHP, ASP, etc., y sobre todo la tecnología *Servlet* de la plataforma J2EE (Java 2 *Enterprise Edition*). Los *servlets* son una interfaz genérica para el desarrollo de aplicaciones servidoras cuya finalidad es la de ofrecer un sistema de servidor dinámico que gestione modularmente el procesado de peticiones HTTP definiendo unas interfaces capaces de independizar la parte de aplicación con la de transporte de la información.

Hoy en día el amplio despliegue de las redes de acceso de banda ancha sumada a la elevada optimización de los algoritmos de compresión y transmisión, abre la posibilidad de ofertar servicios de voz y video en tiempo real a través de redes de conmutación de paquetes, redes originalmente no orientadas a dicho fin. En este entorno se desarrollan las soluciones de telefonía sobre IP (VoIP, Voz sobre IP). En el proceso de desarrollo de esta tecnología surgen nuevos horizontes a medida que los recursos de la red van incrementándose y que pueden ofrecer algo más que el servicio básico de telefonía. De cara a que un operador pueda ofrecer servicios inteligentes de telefonía sobre IP y aprovechando la experiencia y el éxito de los *servlets* en el entorno Web, aparece la implementación de esa interfaz genérica para el protocolo SIP [1] (*Session Initiation Protocol*) utilizado para el transporte de la señalización de las comunicaciones de audio y vídeo.

El resto del artículo se organiza como se indica a continuación. En la Sección 2 se explica brevemente el protocolo SIP. En la siguiente sección se expone la arquitectura básica necesaria para ofrecer servicios avanzados de telefonía utilizando Servlets SIP y que posibilita la convergencia entre aplicaciones SIP y aplicaciones HTTP. En la Sección 3, presentamos nuestra implementación de un servidor de aplicaciones SIP basado en Tomcat, discutiendo su diseño y mostrando las pruebas de carga realizadas. Por último, en la Sección 4 se presenta una

[1] Este trabajo ha sido financiado por el Ministerio de Ciencia y Tecnología de España bajo el proyecto TIC2003-08129-C02.

framework que facilita el desarrollo de servicios en nuestra plataforma con un ejemplo de su uso.

2 El protocolo SIP

SIP (*Session Initiation Protocol*) es un protocolo estandarizado por el IETF (*Internet Engineering Task Force*) utilizado para el transporte de la señalización de comunicaciones de voz y vídeo a través de una red de conmutación de paquetes. Utiliza unos patrones para definir una lógica de intercambio de mensajes de cara a realizar una serie de funcionalidades que se consideran importantes para posibilitar que dos usuarios en Internet sean capaces de transmitir y recibir flujos de voz y vídeo.

SIP proporciona funcionalidades para la determinación de la localización del destino, es decir, resolución de direcciones, mapeo de nombres y redirección de llamadas, la determinación de las capacidades del terminal destino vía el protocolo Session Description Protocol (SDP [2]); la determinación de la disponibilidad del terminal destino; el establececimiento de una sesión entre el terminal original y el terminal destino y para el manejo de la finalización de una llamada o su transferencia a otro terminal.

SIP es un protocolo peer-to-peer (P2P). Los peers que participan en una sesión SIP se denominan User Agents (UA). Un UA puede funcionar indistintamente como cliente o servidor distinguiéndose entonces entre User Agent Client (UAC), que correspondería a una aplicación cliente que inicia una petición SIP, y User Agent Server (UAS), que sería una aplicación servidora que contacta al usuario cuando se recibe una petición SIP y que responde según el comportamiento del usuario. El comportamiento como UAC o UAS de un terminal en una transacción SIP depende, básicamente, de cuál fue el usuario que inició la transacción.

2.1 Arquitectura SIP

Desde el punto de vista arquitectónico una red SIP se compone de elementos que pueden agruparse en dos categorías: clientes y servidores. Los clientes SIP pueden actuar tanto como UAC como UAS e incluyen a los terminales SIP (por ejemplo un teléfono VoIP en el caso de telefonía IP) y a los gateways SIP, dispositivos que proporcionan servicios tales como traducción de formatos y procedimientos de la comunicación entre terminales SIP y otro tipo de terminales no SIP, o transducción entre codecs de audio y vídeo.

Los servidores SIP incluyen:

- Servidores proxy: reciben mensajes SIP y lo reenvía al siguiente servidor SIP en la red. Pueden realizar funciones de autenticación, de autorización, de control de acceso, de encaminamiento, de petición de retransmisiones fiables y de seguridad.

- Servidores de redirección: proporcionan al cliente información acerca del siguiente salto o saltos que un mensaje debería seguir y, entonces, el cliente contacta con el siguiente salto que puede ser tanto otro servidor como el UAS directamente.

- Servidor de registro: procesa las peticiones de los UAC para que registre su localización actual. Por lo general, los servidores de registros actúan coordinadamente con un servidor proxy o de redirección.

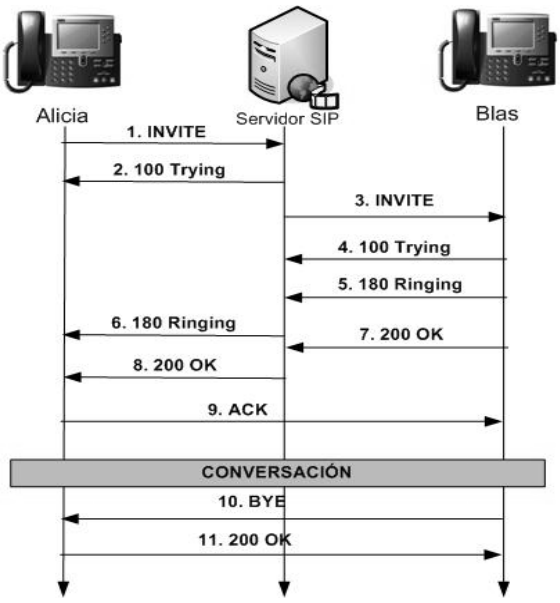


Figura 1 – Establecimiento de llamada en SIP mediante un servidor SIP actuando de Proxy

2.2 Funcionamiento básico de SIP

Los usuarios de una red SIP se identifican por direcciones SIP únicas. Una dirección SIP es similar a las direcciones utilizadas en el correo electrónico y siguen el formato sip:identificador@dominio. El identificador de usuario puede ser tanto un nombre de usuario como, directamente, una dirección IP. Los usuarios se registran en un servidor de registro utilizando la dirección SIP que se les ha asignado. El servidor de registro proporciona esta información al servidor de localización cuando éste la solicite,

Cuando un usuario inicia una llamada, envía una petición SIP a un servidor SIP, ya sea a un proxy o a uno de redirección. Esta petición incluye la dirección SIP del origen de la llamada y la dirección SIP del destino. Ambos se comunican con el servidor de localización para encontrar al usuario SIP final destino. El servidor de localización puede utilizar uno o más protocolos, como, por ejemplo, finger, rwhois y LDAP, para la localización

La Fig. 1 muestra el establecimiento de llamada en SIP utilizando un servidor proxy. El UAC (Alicia) envía un mensaje SIP del tipo INVITE al servidor proxy. El servidor determina el camino al UAS destino (Blas) y le reenvía a éste la petición. El destino responde entonces al servidor proxy, el cual, a la vez, reenvía la respuesta al origen de la llamada. Se establece entonces una sesión entre origen y destino utilizándose el protocolo Real-time Transport Protocol (RTP) para la comunicación entre ambos. Si se utiliza un servidor de redirección, el UAC le envía el INVITE y éste contacta con el servidor de localización para determinar el camino hasta el destino, enviándole esta información al UAC que la utiliza para enviar la petición al dispositivo indicado en la información devuelta y que puede ser tanto un servidor proxy como, directamente, el destino. Al igual que el primer caso, la sesión queda entonces establecida y ambos extremos utilizan RTP para comunicarse.

3 Servidores de Aplicaciones SIP

La funcionalidad de los servidores SIP descrita en la sección 2 por un lado es estática, es decir, realizan las mismas acciones para todas las llamadas y, por otro lado, se limitan a los servicios ya descritos. Los servidores de aplicaciones son componentes encargados de la ejecución de aplicaciones externas que añaden inteligencia a los servicios que puede ofrecer una red SIP. Se entiende por servidor de aplicaciones a la entidad capaz de resolver peticiones dinámicamente a través de pequeños programas que se ejecutan al servidor y que responden en función de los parámetros de dichas peticiones. Actualmente los servidores de aplicaciones en el mundo Web están ampliamente desplegados. La idea de asimilar esa solución Web al entrono SIP busca ofrecer acciones dinámicas sobre determinados eventos que se pueden producir en el establecimiento o durante el transcurso de una conversación telefónica, aunque, como se verá, no sólo se limita a la Telefonía IP, sino que se puede extender como solución de red para la oferta de todo tipo de servicios multimedia y peer-to-peer.

La solución que se ha adoptado como Servidor de Aplicaciones SIP se basa en un contenedor de servlets adaptado al protocolo SIP, en adelante, servlets SIP. Los servlets SIP están estandarizados en una Java Specification Request (JSR[3]). Una JSR es un proceso de estandarización de la comunidad Java que consiste en plasmar una idea para una especificación Java de cara a ser evaluado por la comunidad Java (Java Community Process, JCP [3]) de modo similar al que sigue el IETF con los Internet Drafts y los Request For Comments (RFC). La especificación "SIP Servlet API [4]", se encuentra en estado "Final", lo cuál significa que se dispone de un documento de libre consulta de cara a que los desarrolladores interesados puedan implementar dicha especificación.

La "SIP Servlet API" detalla el entorno de funcionamiento general del entorno, las clases y métodos que deben implementarse, y el modo en el que se relaciona el contenedor con los servlets SIP y el entorno. Todo proceso JSR que alcance el estado "Final" proporciona, además, una implementación de referencia que tiene como objetivo dar un ejemplo de funcionamiento y composición conceptual del sistema para facilitar tanto la comprensión de la especificación como la creación de una implementación base que sirva como entrono de pruebas para verificar el cumplimiento de la especificación. La implementación de referencia de una JSR no suele ser completamente funcional y no debe ser usada en entornos reales de producción.

3.1 JAIN SIP

JAIN [5] es un conjunto de APIs (*Application Programming Interface*) que tiene como función principal agrupar e implementar interfaces necesarias para el desarrollo de nuevos servicios en el ámbito de las telecomunicaciones.

Dentro de la comunidad de JAIN, se encuentra JAIN SIP que es la parte encargada de implementar las interfaces, para la utilización del protocolo SIP. La especificación de JAIN SIP proporciona funcionalidades definidas en el RFC 3261 (SIP) y en los RFCs que complementan al protocolo SIP, alguna de las funcionalidades son: Métodos para el formato de los mensajes , capacidad a la aplicación para poder recibir y enviar mensajes, capacidad para poder analizar los mensajes recibidos y acceder a sus campos, transporte de información de control generada en una sesión (RFC 2976), mensajería instantánea (RFC 3428). Otros RFCs que complementan la especificación de SIP son el 3262, 3265, 3311, 3326 y 3515.

3.2 Servlets SIP

Al igual que los servlets HTTP, un servlet SIP tiene un funcionamiento orientado a eventos, de manera que se puede definir su comportamiento en función del mensaje SIP recibido. Pero a diferencia de los servlets HTTP, los servlets SIP pueden recibir y generar tanto peticiones como respuestas (PUSH). Esta diferencia es debida a que SIP es un protocolo peer-to-peer mientras que HTTP esta basado en la arquitectura cliente/servidor.

La finalidad de un servlet SIP consiste en posibilitar la programación tanto de los servicios fundamentales de un servidor SIP como la de servicios de valor añadido que superan ampliamente los requerimientos funcionales de un servidor SIP. Ejemplos de estos últimos podemos citar servicios de presencia, de mensajería instantánea ante la no posibilidad de atender una llamada, centralita de telefonía automática, buzón de voz, redirección de llamadas, etc.

Para implementar un servlet SIP, el programador parte de una interfaz definida por la especificación y cuyo diagrama UML se muestra en la Fig. 3, dónde se muestran los métodos a implementar por todo servlet SIP. Cada uno de los métodos doX() se invoca cuando el servlet recibe el mensaje SIP "X" correspondiente. Así, por ejemplo, la implementación del método doRegister en un servlet se invocaría cuando el servidor recibe un mensaje SIP REGISTER, y el código a realizar correspondería al registro del usuario que ha enviado el mensaje dentro del sistema de localización de la red SIP.

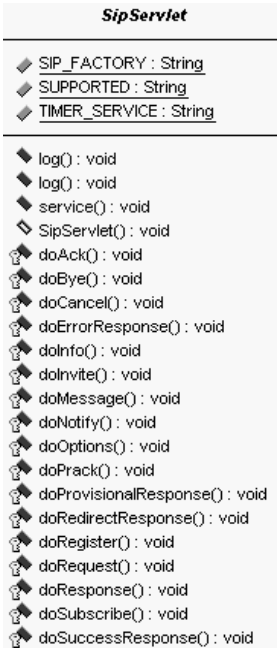


Figura 2 – Interfaz de los Servlets SIP

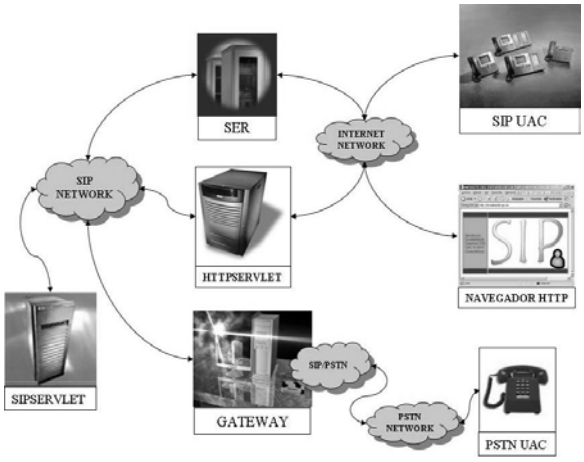


Figura 3 – Arquitectura de una Red SIP

3.3 Componentes de una Arquitectura con Servidor de Aplicaciones SIP

La Fig. 3 muestra los componentes básicos de una red SIP que utiliza un servidor de aplicaciones SIP para ofrecer tanto servicios básicos de Telefonía IP con conectividad contra la Red Telefónica Conmutada como servicios avanzados utilizando un servidor de aplicaciones SIP y un servidor de aplicaciones Web.

Esta arquitectura utiliza un servidor SIP, como podrían ser SER o VOCAL, por citar dos ejemplos de servidores SIP de libre distribución. Toda la funcionalidad de este servidor, ya descrita en la Sección 2, podría ser implementada dentro del servidor de aplicaciones SIP por uno o varios Servlets. De hecho se podría partir de la arquitectura presentada en la Fig. 3 e ir gradualmente migrando hacia el uso único de un servidor de aplicaciones SIP, ya que este corresponde a un conjunto de servicios que incluye los servicios básicos de un servidor SIP. El componente gateway es necesario para poder comunicar terminales IP con terminales de redes de telefonía conmutadas.

El servidor de aplicaciones Web posibilita la integración de aplicaciones Web en la red SIP. Por ejemplo, considérese un usuario con un terminal que dispone de la capacidad de realizar llamadas y un navegador Web. Dicho usuario inicia una llamada contra otro usuario de la red que, por alguna razón, en ese momento no puede atender la llamada. El servlet SIP recoge este evento, se lo notifica al usuario y le redirecciona hacia una aplicación Web en la que puede escribir un mensaje corto de texto que le será enviado al usuario con el que intentaba establecer comunicación. En [6] se presenta otro ejemplo de aplicación que explota las posibilidades de convergencia entre una red SIP y las aplicaciones Web.

3.4 Implementación de un Servidor de Aplicaciones SIP

A partir de la implementación de referencia se ha desarrollado un servidor de aplicaciones SIP en convenio (UPC-CTT C05359) con la empresa Voztelecom S.L (http://www.voztele.com). Se ha desarrollado como un componente para Tomcat [7], un contenedor de Servlets que se ha convertido en la implementación de referencia para las tecnologías Java Servlet y Java Server Pages. Para ello se han implementado tanto el conector Tomcat que permite levantar un proceso que escuche en un puerto y reciba los mensajes SIP, como el procesador Tomcat que se encarga de instanciar como objeto java SipRequest los octetos recibidos. Otra función importante que realiza este módulo consiste en hacer llegar ese objeto hasta la aplicación correspondiente. Esto se consigue a través de unas reglas que define el usuario.

La ventaja fundamental de integrar el contenedor de los SIP Servlets en Tomcat es que permite, entre otras cosas, la posibilidad de almacenar información de sesión multiprotocolo, en este caso HTTP y SIP, facilitando al programador de Servlets la creación de aplicaciones en las que interactúen accesos Web con llamadas VoIP. Así por ejemplo, se podría implementar una aplicación WEB en la se

monitorizase los estados de los teléfonos de una oficina (llamada en curso, libre, iniciando, pendientes, etc.) e incluso, creando hiperenlaces sobre los número de teléfono provocar del inicio de una llamada, contestar a las pendientes, cancelarlas, redirigirlas, etc.

A continuación destacamos las características fundamentales del servidor de aplicaciones SIP desarrollado:

- Permite el uso tanto de UDP como de TCP para el transporte de SIP.

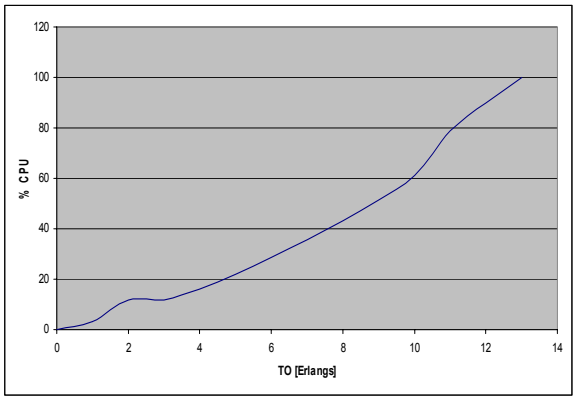


Figura 4 – Consumo de CPU vs. Carga

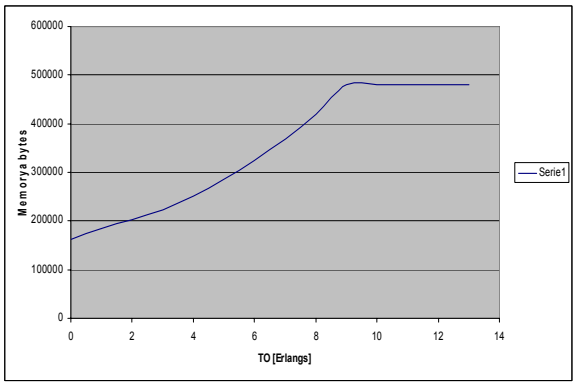


Figura 5 - Consumo de Memoria vs. Carga

- Permite el desarrollo de aplicaciones convergentes HTTP/SIP.
- Utiliza la JAIN SIP para el procesado de mensajes SIP en java.
- Implementa un servidor SIP completo, es decir, que puede actuar de proxy, de servidor de registro y de redirección.
- Posee una API que implementa MSML/MOML [8], dos protocolos que facilitan la comunicación con servidores multimedia.
- Incorpora una *framework* de desarrollo de aplicaciones que se explica en la sección 4.

3.5 Pruebas de carga del Servidor de Aplicaciones

En esta sección presentamos las pruebas de carga que se han realizado sobre la actual implementación del servidor de aplicaciones SIP que hemos desarrollado. Un UAC inicia una sesión enviando un INVITE al servidor de aplicaciones SIP que, actuando como proxy, lo reenvía directamente al UAS. La llamada tiene una duración media de 1 segundo, según un modelo de servicio exponencial. Transcurrido ese tiempo, el UAC finaliza la llamada enviando un BYE.

El servidor se ejecuta sobre un PC con procesador Intel Premium M a 1,6GHz con 500 Mbytes de memoria. El sistema operativo es UNBUTU Debian con un kernel Linux 2.6.8.1-3-386 con una memoria swap disponible de 530Mbytes. El sistema arranca en modo 3 para conseguir que el sistema operativo arranque los mínimos procesos necesarios para su funcionamiento y que interfiera lo mínimo posible en las medidas obtenidas.

Las pruebas realizadas consisten en ir aumentando cada minuto el tráfico ofrecido desde 1 Erlang hasta 13 Erlangs. La Fig. 4 y la Fig. 5 muestran los resultados de consumo de CPU y de memoria, respectivamente, obtenidos. Como se observa, la implementación permite una carga máxima, bajo esas condiciones de tráfico, de 10 Erlangs. Teniendo en cuenta las limitaciones propias de la máquina donde se han desarrollado las pruebas nos hace albergar grandes expectativas cuando se disponga de una máquina especialmente preparada para actuar como servidor.

4 Servicios SIP Multimedia

La utilización clásica del protocolo SIP ha sido como protocolo de señalización para Telefonía IP. Sin embargo el rango de aplicación de SIP puede extenderse a la señalización para facilitar el intercambio de todo tipo de información multimedia entre usuarios como imágenes, videos o juegos en red multijugador, tanto para red fija como para redes móviles, e intentando mantener un alto grado de filosofía P2P en la transmisión. En esta sección presentamos una framework que permite el desarrollo de servicios sobre un servidor de aplicaciones SIP de modo sencillo y que permite un alto grado de reutilización así como un servicio SIP que estamos desarrollando sobre dicha framework a modo de ejemplo.

4.1 Framework SIP Servlets

Se ha diseñado una framework para el desarrollo de servicios basados en SIP. En este diseño se ha seguido un modelo arquitectónico basado en 3 capas.

En primer lugar se tiene el servidor de aplicaciones SIP (SIP AppServer) que implementa una pila SIP (basado en JAIN-NIST), un contenedor de servlets SIP, gestión de las sesiones SIP, temporizadores, etc. La siguiente capa se divide en dos partes. Por un lado se tiene el CallControl, que es un elemento encargado del control de las llamadas y la redirección de la llamada al servicio correspondiente. La implementación consiste en un único SipServlet que gestiona los servicios implementados. Se ha desarrollado siguiendo el patrón de diseño Front Controller (Core J2EE Patterns [9]) y respetando la especificación JAIN™ Java Call Control (JCC-JSR21). La segunda parte consiste en los componentes, es decir, el conjunto de objetos que facilitarán el diseño e implementación de los servicios. Proporcionan un nivel de abstracción para la implementación de los servicios: RTPPRoxy (API de control sobre RTPProxy de SER que permite la implementación de soluciones compatibles a NAT's y cortafuegos), MediaServer, un componenente que modela un servidor de contenidos multimedia, y un SIPActor, componente que modela un UAC y/o UAS dentro de la lógica de un servicio.

Por encima de este middleware se ejecutan los servicios. Para minimizar el desarrollo de los servicios, el *time-to-market* y de cara a facilitar la implementación de los servicios se propone un desarrollo orientado a maquinas de estados (patrón de diseño [10]) . Una máquina de estados esta formada por de estados y transiciones. Los estados representan el momento actual de la máquina, cuando recibe un entrada (*input*) se ejecuta la lógica asociada a una transición y la máquina puede o no cambiar de estado. También se pueden definir condiciones que se evalúan cuando se recibe una entrada, de manera que una misma entrada puede conducir a un estado u otro dependiendo del resultado de la condición.

4.2 Ejemplo

A continuación se presenta a modo de ejemplo un servicio implementado sobre el *framework* explicado anteriormente que ilustra un servicio que se beneficia de las ventajas de un servidor de aplicaciones convergente WEB - SIP. Se trata de un servicio de conferencia que ofrece dos funcionalidades: planificación de la conferencia y puesta en marcha de la misma.

En la Fig. 6 se presenta un ejemplo de planificación de conferencia para el 8 de Marzo de 2005, entre las 17:00 y las 17:30, cuyos participantes son Sergio, Xavier y Antonio

Figura 6 – Interfaz Web de la aplicación de planificación de conferencia.

La segunda parte del servicio tiene como requisito previo que exista una conferencia planificada previamente en el sistema. La conferencia tiene asignada una fecha, hora y participantes de la misma. La referencia sobre los participantes se realiza usando las direcciones SIP (SipUri) de los participantes

En el momento que coinciden día y hora del sistema con día y hora de la conferencia planificada se habilita el proceso de configuración y puesta en marcha de la conferencia.

El sistema inicia el proceso de configuración y puesta en marcha de la conferencia de manera automática. Se crea una sala virtual que acogerá a los participantes y se queda en un estado pendiente de configuración. La configuración se completa en el momento que se ha *invitado* a los participantes de la conferencia y éstos han confirmado (no tienen por que hacerlo todos). En la Fig 7 se presenta la máquina de estados que modela esta segunda parte del servicio.

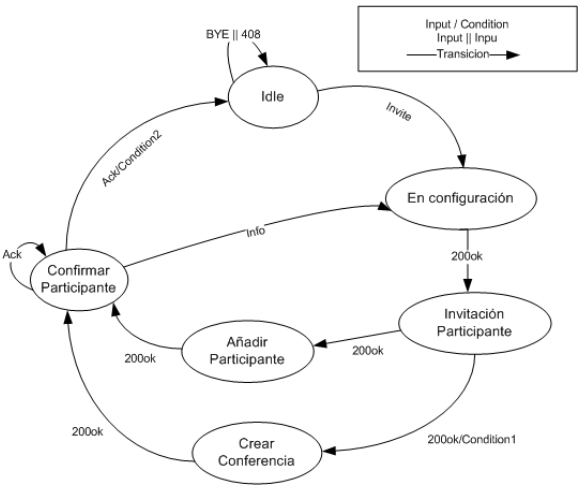


Figura 7 - Máquina de estados de un servicio de conferencia (segunda fase).

5 Conclusiones

En este artículo hemos presentado una arquitectura para una red SIP basada en el uso de un Servidor de Aplicaciones SIP que permite el desarrollo de servicios avanzados para este tipo de redes. Se han

presentado las pruebas de carga realizadas sobre una implementación de la especificación de los Servlets SIP integrada en el servidor de aplicaciones Tomcat, lo cual permite una convergencia entre aplicaciones SIP y HTTP. También se ha explicado las líneas básicas de diseño de una *framework* para el desarrollo de dichos servicios, así como un ejemplo de uso de la misma.

Referencias

[1] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: Session Initiation Protocol," IETF RFC, 2002.

[2] M. Handley and V. Jacobson, "SDP: Session Description Protocol," IETF RFC, 1998.

[3] S. MicroSystems, "JCP: Java Community Process," http://www.jcp.org/.

[4] S. MicroSystems, DynamicSoft, "JSR Sip Servlet API," http://www.jcp.org/aboutJava/communityprocess/final/jsr116/, 2003.

[5] S. MicroSystems, *Jain API*: http://java.sun.com/products/jain/.

[6] Xavier Hesselbach, Sergio Machado, Antonio Abajo, "Servicios Avanzados de Videoconferencia," presented at XIV Jornadas Telecom I+D, Madrid, Barcelona, Valencia y Malaga, Noviembre 2004.

[7] A. S. Foundation, *Apache Jakarta Tomcat*: http://jakarta.apache.org/tomcat/.

[8] C. Systems, "MSML / MOML Specification," http://www.ietf.org/internet-drafts/draft-even-media-server-req-00.txt.

[9] S. MicroSystems, *Core J2EE Patterns*: http://java.sun.com/blueprints/corej2eepatterns/.

[10] B. Eckel, "Thinking in Patterns, Problem-Solving Techniques using Java," http://www.mindview.net/Books/TIPatterns/, 1999.

Malaca: Una Arquitectura para el Desarrollo de Agentes Software basada en Componentes y Aspectos

Mercedes Amor y Lidia Fuentes
Departamento de Lenguajes y Ciencias de la Computación. Universidad de Málaga
ETSI de Informática. Campus de Teatinos S/N.
29071 – Málaga
Teléfono: 952 13 27 96 Fax: 952 13 13 97
E-mail: {pinilla, lff}@lcc.uma.es

Abstract. *Current agent architectures implemented by MAS platforms normally lead the developer to implement the domain-specific functionality and interaction of software agents from scratch, placing little emphasis on (re)configuration and (re)use. This paper presents Malaca, a component and aspect–based software agent architecture that promotes building agents from reusable software components and the configuration of some software aspects. The basis of our architecture is the use of component-based and aspect-based software development concepts to separate agent functionality and crosscutting concerns into independent entities increasing extensibility, maintainability and adaptability of the agent to new environments and demands. The architecture simplifies the software agent development process, which can be reduced to the description of the agents´ constituent components and supported agent interaction protocols using XML documents.*

1 Introducción

Para extender el uso de la tecnología de agentes es necesario, entre otras cosas, mejorar el proceso de desarrollo de sistemas multi-agente (SMA) definidos por plataformas actuales. En la actualidad, la construcción de agentes sobre cualquiera de las soluciones existentes resulta, en la mayoría de los casos, una tarea compleja y tediosa, propensa a errores, y requiere que el desarrollador tenga conocimientos en un lenguaje de programación concreto, normalmente uno orientado a objetos. Esto significa que, una vez que el desarrollador adquiere los conocimientos necesarios para programar un agente sobre una plataforma de agentes concreta, no se interese por el uso de otras APIs y marcos de trabajo proporcionados por otras plataformas, aunque estas resulten más adecuadas que la que conoce. Nuestro objetivo es doble, por un lado facilitar el trabajo del desarrollador simplificando la tarea de programar agentes software, y por otro, fomentar la (re)utilización de agentes dentro de diferentes plataformas de agentes.

Esencialmente las arquitecturas de agentes actuales se diseñan e implementan como marcos de trabajo orientados a objetos (OO) [1,2,3] que proporcionan una colección extensible de clases e interfaces que modelan conceptos típicos de los agentes. El Desarrollo de Software basado en Componentes (DSBC) [4] surge como una evolución del desarrollo de software orientado a objetos, y dota a los lenguajes orientados a objetos de capacidades de composición más allá de la herencia. Así, mientras que los objetos se expresan a nivel de lenguaje, los componentes se expresan principalmente explorando su interfaz pública y fomentando la reutilización de *caja negra*[5].

Por tanto sería muy beneficioso aplicar para la construcción de SMA el disponer de sistemas de desarrollo basados en componentes. Nuestra propuesta se basa en definir una arquitectura basada en componentes para el desarrollo de agentes software llamada Malaca. Aplicando el DSBC descomponemos la funcionalidad del agente dependiente de un dominio de aplicación concreto en componentes software independientes, que pueden ser componentes propios, componentes COTS (del inglés *Commercial Off-The-Shelf*), o incluso Servicios Web[6]. De esta forma la construcción de agentes software se realiza mediante el ensamblado de componentes COTS reutilizables, reduciendo los tiempos, costes y esfuerzos requeridos durante el desarrollo. Al usar componentes COTS ya probados se minimiza la aparición de errores y el tiempo dedicado a la depuración. Además, los agentes software resultantes son más adaptables, flexibles y reutilizables y se evita su desarrollo desde cero.

Sin embargo, la descomposición de un sistema, y en particular de los sistemas complejos como los agentes software en componentes independientes, no es una tarea trivial. Normalmente ocurre que un mismo concepto o propiedad suele estar presente en varios componentes del sistema, creando dependencias entre ellos. Como consecuencia estas propiedades se "entremezclan" con la funcionalidad afectando a la reutilización y evolución del componente software. Esto se conoce como el *problema del código enmarañado*[7]. En este sentido, la separación de conceptos avanzada y más concretamente el Desarrollo de Software Orientado a Aspectos (DSOA) [8] propone identificar y modelar estas

propiedades extra-funcionales separadas de la funcionalidad. El DSOA (en terminología inglesa AOSD, *Aspect Oriented Software Development*) es una disciplina muy prometedora cuyo objetivo principal es mejorar el proceso de desarrollo de software, en un intento de superar la creciente complejidad de los sistemas software. Siguiendo el enfoque propuesto por el DSOA las propiedades extra-funcionales de los agentes se modelan como *aspectos*, considerados entidades de primer orden. De esta forma en la arquitectura interna de un agente Malaca, propiedades como la coordinación entre agentes y la distribución de la comunicación se modelan como entidades diferentes y desacopladas de los componentes, encargados de proporcionar la funcionalidad. La composición de estos aspectos y de los componentes internos del agente se lleva a cabo en tiempo de ejecución, permitiendo la (re)configuración y adaptación del comportamiento del agente para incorporar, por ejemplo, nuevos protocolos de coordinación y nueva funcionalidad.

El hecho de que en los agentes Malaca la coordinación se modele e implemente de forma separada a la funcionalidad del agente mejora de forma importante su desarrollo, uno de los objetivos de este trabajo. Al desacoplar la funcionalidad del agente de la interacción en la que participa, se facilita la (re)utilización tanto de la funcionalidad como de la coordinación del agente durante el proceso de desarrollo del agente. De igual modo, otras propiedades del agente que dependen de la plataforma sobre la que se ejecutará se encapsulan en aspectos independientes. Así, el uso de un servicio de transporte para distribución de mensajes es encapsulado en el aspecto de distribución. Análogamente, la codificación de mensajes en una representación FIPA también es representada en un aspecto de representación. Dado que las dependencias de la plataforma son encapsuladas en aspectos independientes, los agentes Malaca pueden ser adaptados para ser ejecutados sobre diferentes plataformas de agentes, consiguiendo el segundo de los objetivos de este trabajo. De esta forma, tan sólo es necesario programar los agentes una vez para utilizarlos sobre diferentes plataformas de agentes FIPA.

Otra importante contribución de nuestra propuesta respecto a la mejora del desarrollo de agentes software es que un agente puede ser "programado" simplemente editando documentos XML que describan y configuren los componentes que constituyen su arquitectura interna. De esta forma, el desarrollador sólo debe proporcionar una descripción de los componentes que serán ensamblados en la arquitectura del agente, una descripción explícita de los protocolos de interacción soportados, e información acerca de la configuración y despliegue de otros aspectos del agente.

Para mostrar los beneficios de nuestra propuesta presentaremos a modo de ejemplo en la sección 2 el desarrollo de un agente vendedor de libros en JADE, uno de los marcos de trabajo de agentes más usados. Posteriormente, en la sección 3, mostraremos el mismo ejemplo desarrollado en Malaca y en la sección 4 ilustraremos mediante una pequeña comparativa las bondades y beneficios que proporciona el uso de nuestra propuesta en relación al desarrollo de agentes software.

2 Desarrollo de Agentes Software en JADE

En esta sección describimos cómo construir un agente software utilizando JADE, uno de los marcos de trabajo más populares para la construcción de agentes FIPA. Tomaremos como caso de estudio un agente vendedor de libros cuya implementación se incluye en el ejemplo *bookTrading*, disponible en la distribución de JADE[3].

2.1 JADE

JADE (Java Agent DEvelopment Framework)[3] ofrece una plataforma para la ejecución de agentes software que implementa las directrices de FIPA y define un marco de trabajo de desarrollo de agentes implementado sobre Java. Este marco de trabajo proporciona un conjunto de clases e interfaces sobre los cuales implementar la funcionalidad del agente y su comunicación dentro de una aplicación concreta. Cabe destacar que JADE es Software Libre.

Una de las principales características de esta plataforma es que cumple las especificaciones definidas por el estándar FIPA[9]. El intercambio de mensajes entre agentes, así como las *performativas* empleadas se corresponderán con lo especificado con este estándar. De igual forma, los servicios que debe de proporcionar la plataforma para el transporte de mensajes (Servicio de Transporte de Mensajes o MTS), el servicio de gestión de agentes (AMS) y el servicio de directorio (DF) también siguen las especificaciones dadas por FIPA a ese respecto. La plataforma de ejecución de agentes se caracteriza por permitir la ejecución de diversos agentes en varias plataformas JADE de forma concurrente, e incluso facilita el desplazamiento de estos agentes de una máquina a otra.

Respecto a la programación de agentes, para construir un agente sobre el marco de trabajo definido por JADE es necesario implementar un conjunto de clases Java que extiendan las ya definidas en el marco de trabajo. En concreto, será necesario crear, como mínimo, una clase Java que derive de la clase *jade.core.Agent* por cada tipo de agente definido en el sistema multi-agente (fig.1). Esta clase deberá incluir básicamente un método *setup()*, que se ejecuta al iniciarse el agente y que contendrá código de inicialización, incluyendo la especificación e instanciación de los comportamientos asociados al agente. Cada agente definirá en diferentes clases Java

derivadas de la clase *jade.core.behaviours.Behaviour* cada uno de los comportamientos asociados al agente. El comportamiento de un agente JADE define la funcionalidad específica del agente en una aplicación, e incluye código referente tanto a la funcionalidad como a su interacción dentro de un sistema multi-agente específico. Básicamente estos comportamientos determinan el envío y recepción de mensajes, relacionados asimismo con la realización de tareas y funciones propias del dominio de aplicación del agente. Cuando el agente es creado, o en tiempo de ejecución, estas clases son añadidas al conjunto de comportamientos del agente mediante el método *addBehaviour*() de la clase *Agent*.

2.2 Implementación de un Agente Vendedor de Libros en JADE

La Fig.1 muestra, mediante un diagrama de clases UML, la implementación del agente vendedor de libros en JADE. Las clases sombreadas en gris representan las clases que debe implementar el programador. Las clases sin sombrear revelan qué clases son definidas en el marco de trabajo proporcionado por JADE debiendo ser extendidas para construir el agente. El método *setup()* de la clase *BookSellerAgent*, que extiende a la clase *Agent* de JADE. Como parte de este método, que es invocado cuando el agente es creado, se realiza el registro del agente en el agente DF de la plataforma (servicio de directorio), y se crean e inicializan los recursos y comportamientos del agente vendedor de libros.

El comportamiento del agente se modela mediante las clases *OfferRequestsServer*, *PurchaseOrdersServer* y *UpdateCatalogue*. El comportamiento encapsulado en la primera se encarga de atender peticiones enviadas por agentes que compran libros. Estas peticiones, enviadas en mensajes ACL del tipo CFP (*call-for-proposal*), incluyen el título del libro buscado. El agente vendedor toma del contenido el título del libro y consulta si figura entre los incluidos en su catálogo. El catálogo del agente contiene los títulos de los libros y su precio. Si el agente vendedor dispone de dicho libro entre sus existencias envía al agente comprador un mensaje ACL del tipo *propose* que contiene el precio del libro. En caso contrario, le notifica, mediante un mensaje ACL de tipo *refuse,* que no dispone de dicho título. El desarrollador implementará este protocolo de interacción en el método *action()*.

El comportamiento definido en la clase *PurchaseOrdersServer* es el encargado de llevar a cabo la venta de un libro. Este comportamiento gestiona los mensajes ACL *accept-proposal* enviados por los agentes compradores en respuesta a una propuesta acerca de la venta de un libro. Como efecto de la realización de dicha venta, el agente eliminará dicho título del catálogo e informará al agente comprador de que se ha efectuado la compra. Si al intentar eliminar el libro del catálogo éste ya no estuviera, significa que mientras se realizaba la interacción el libro ya ha sido vendido a otro agente. En este caso el agente vendedor comunicará al agente comprador que la venta no ha podido realizarse mediante un mensaje del tipo *failure*. Este protocolo de interacción también se implementa en el método *action()* de la clase *PurchaseOrdersServer*. Los comportamientos de ambas clases extienden del comportamiento cíclico de JADE, lo que significa que estarán siempre activos durante la ejecución del agente. El tercer comportamiento, encapsulado en la clase *UpdateCatalogue* se encarga de actualizar el catálogo de libros. Esta actualización es realizada por el usuario a través de una interfaz gráfica modelada por la clase *BookSellerGUI*, y se ejecutará cada vez que el usuario añada un nuevo libro al catálogo. La ejecución del método *action(),* donde se incluye el código relativo a la actualización del catálogo, se ejecutará de forma atómica por lo que extiende de la clase *OneShotBehaviour*.

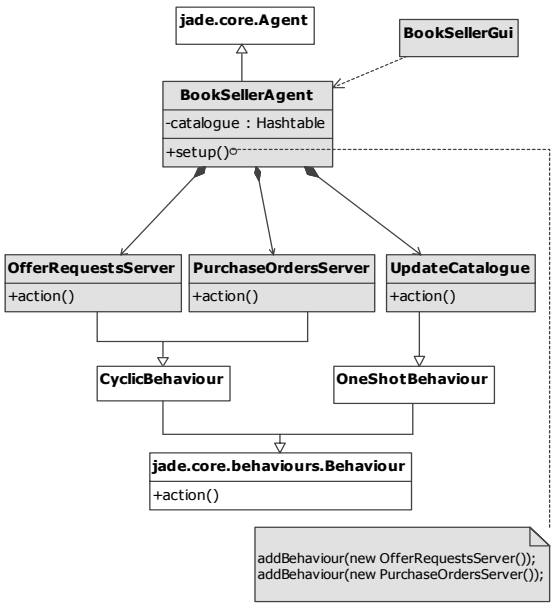


Figura 1. Diagrama de clases UML de un Agente Vendedor de Libros desarrollado en JADE.

Cabe destacar que en las clases *OfferRequestsServer* y *PurchaseOrdersServer* los protocolos de petición y compra de libros se encuentran entremezclados con el código de acceso al catálogo. Esto hace que se dificulte la reutilización en otros sistemas del mismo dominio, tanto de la funcionalidad específica del catálogo como de los protocolos de interacción. Por otro lado, todas las clases que implementan la funcionalidad específica en este agente JADE acceden al catálogo de libros, por lo que podemos considerar que hay una dependencia no deseable de cara a que se modifique la interfaz de acceso a dicho catálogo. El desarrollador de agentes utilizando Malaca representa una mejora considerable a estas limitaciones proporcionando una mayor modularización del agente gracias a la aplicación de los principios del DSBC y del DSOA.

3 Desarrollo de Agentes Software sobre Malaca

En esta sección describiremos brevemente las características más importantes de la arquitectura interna de un agente Malaca y a continuación mostraremos cómo se programaría el agente vendedor de libros utilizando nuestra propuesta.

3.1 Arquitectura de un Agente Malaca

La Fig.2 muestra el diagrama de clases UML de la arquitectura Malaca basada en componentes y aspectos propuesta, instanciada para el ejemplo del agente vendedor de libros. Actualmente, la arquitectura está implementada en Java, y proporciona un marco de trabajo, denominado Malaca, que permite la construcción de agentes software. Al igual que para JADE, en el diagrama de clases de la Fig. 2 se muestran sombreadas en gris los elementos (clases y documentos XML expresados mediante notas) que el programador debe crear para construir el agente vendedor de libros utilizando Malaca. En color blanco se indican las clases y componentes de la arquitectura proporcionados por el marco de trabajo. La arquitectura de un agente Malaca se compone de dos tipos de entidades, los componentes software y los aspectos (representados en la Fig. 2 por el estereotipo *SoftwareComponent* y la clase *Aspect* respectivamente). Los componentes software se encargan de proporcionar la funcionalidad del agente. Algunos componentes están siempre presentes en la arquitectura ofreciendo la funcionalidad básica necesaria para, por ejemplo, crear un mensaje (clase *BasicAgentActions* que representa al componente del mismo nombre). Además de la funcionalidad común, la funcionalidad propia de un dominio de aplicación también residirá en componentes software. Así, para el ejemplo del agente vendedor de libros, la funcionalidad propia de este agente necesaria para comprobar la existencia de un libro en el catálogo o efectuar la venta de un libro, será proporcionada por el componente *BookSellerComponent*. La arquitectura Malaca no impone ninguna restricción acerca de los componentes, por lo que en realidad cualquier componente software puede incorporarse como parte de la funcionalidad del agente, desde un componente COTS (*Commercial Off-The-Shelf*) hasta un servicio Web [6]. Esta flexibilidad se debe a que extendemos el uso de OWL-S [10], una ontología aplicada a la descripción de servicios, para describir la interfaz pública de los componentes que proporcionan la funcionalidad del agente de forma independiente a una implementación. Por tanto para incluir un componente en la arquitectura de un agente tan sólo es necesario proporcionar una descripción de la interfaz pública del componente en XML y OWL-S.

Como comentamos en la introducción, aplicando los principios del DSOA en Malaca modelamos la coordinación de forma separada de la funcionalidad de dominio del agente. En otras propuestas dentro de un mismo componente o clase de la arquitectura de un agente encontramos, además de su comportamiento interno, código relativo al procesamiento y gestión de mensajes de comunicación. Por ejemplo, volviendo al ejemplo del agente comprador de libros implementado en JADE mostrado en la sección anterior, observamos que en la clase *OfferRequestsServer* se concentra el comportamiento del agente orientado a atender las peticiones de libros enviadas por agentes compradores (método *action()* de dicha clase). Dentro de este método se incluye la recepción y el procesamiento del mensaje de petición y la generación del mensaje de respuesta. Este diseño hace difícil reutilizar sólo la parte funcional de esta clase cuando varían aspectos de la comunicación, como por ejemplo el tipo de mensaje (supongamos simplemente que en lugar de recibir un mensaje de tipo *CFP* debiera gestionar mensajes de tipo *request*). Además del aspecto de coordinación también modelamos como aspectos otras propiedades extra-funcionales que describen y caracterizan la comunicación del agente, como la representación de mensajes ACL y la distribución de mensajes a través de una plataforma de agentes concreta.

En la Fig.2 las clases *CoordinationAspect*, *StringRepresentationAspect*, y *DistributionAspect* representan los aspectos de coordinación, representación y distribución respectivamente del agente vendedor de libros tomado como ejemplo. Los aspectos extienden la clase *Aspect* que modela el comportamiento general de un aspecto en el marco de trabajo Malaca. Esta clase define un método *handleMessage()* que será invocado cuando se requiera la aplicación o evaluación del aspecto. Por tanto la funcionalidad del aspecto debe ser implementada en ese método. Los aspectos incluidos en la arquitectura del agente se aplicarán en dos momentos de la ejecución del agente; cuando el agente envía y cuando el agente recibe un mensaje. La aplicación de aspectos al procesamiento de mensajes de entrada y salida del agente se define mediante un conjunto de reglas que determinan cómo deben componerse o "entretejerse" los aspectos cuando deben aplicarse varios a un mismo mensaje. Esta información es utilizada por el componente *Mediador* para llevar a cabo la composición de componentes y aspectos en tiempo de ejecución.

El aspecto de coordinación, representado en la arquitectura por conectores (clase *CoordinationAspect* de la Fig.2) se encarga de coordinar las diferentes interacciones o conversaciones en las que participa el agente de acuerdo a un protocolo de comunicación. Aunque los conectores se encarguen de coordinar la ejecución de un protocolo, los patrones de interacción no se encuentran codificados dentro de los conectores como es usual.

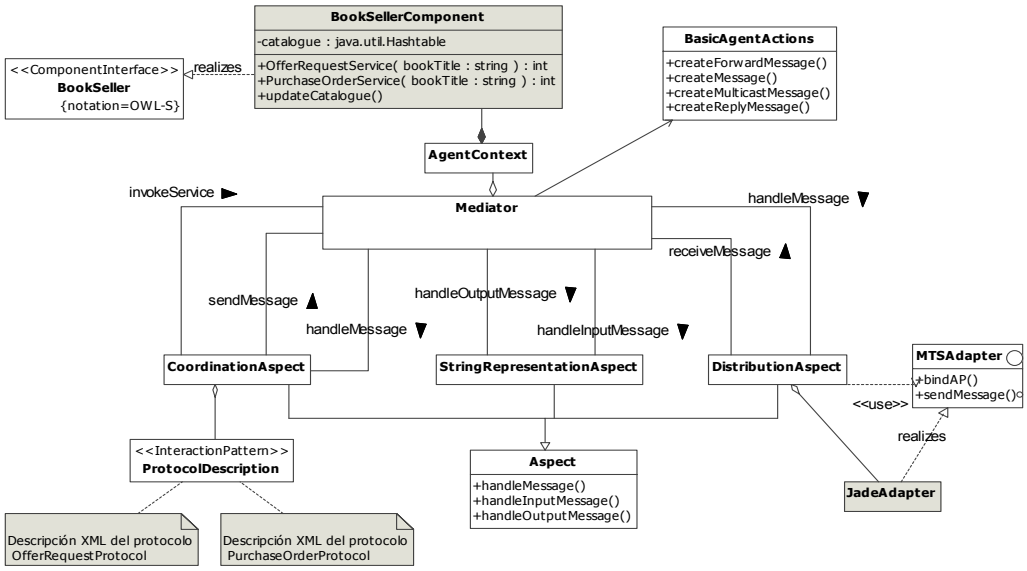


Figura 2. Diagrama de clases UML de la Arquitectura Malaca del Agente Vendedor de Libros.

En su lugar, cuando un conector es creado acepta una descripción de un protocolo de coordinación y controla su ejecución en base a ésta (las descripciones se señalan cómo notas sombreadas en gris en la Fig.2). La descripción de un protocolo define no sólo el intercambio de mensajes que se lleva a cabo, sino también qué acciones internas lleva a cabo el agente durante la ejecución del protocolo. De esta forma es posible enlazar o conectar la funcionalidad de un agente con su participación en una interacción. Un agente puede participar en más de una conversación de forma simultánea, y cada conversación será controlada por una instancia del aspecto de coordinación diferente. Realmente todos los conectores instanciados son iguales, y sólo difieren en el protocolo que coordinan, el cual es suministrado al conector cuando es instanciado. Por tanto, el desarrollador no tiene que implementar un conector para cada protocolo de interacción en que participe el agente. Tan sólo debe proporcionar una descripción del protocolo de interacción en XML. El conector encargado de controlar y coordinar dicha conversación utilizará esta descripción para llevar a cabo su tarea. Veremos con mayor detalle cómo se describe un protocolo en XML en la siguiente sección. Puede encontrar una descripción más detallada del funcionamiento e implementación del conector en [11].

El aspecto de representación, mostrado en la Fig.2 por la clase *StringRepresentationAspect,* se encarga de codificar y decodificar los mensajes de entrada y salida en diferentes representaciones del lenguaje de comunicación de agentes (ACL). En el caso de los mensajes de entrada, este aspecto descarta aquellos que contienen errores sintácticos de acuerdo a su representación. El uso de una representación concreta para codificar los mensajes pertenecientes a la comunicación del agente será configurada como parte de la información de despliegue del agente.

En la arquitectura de un agente Malaca la comunicación con otros agentes o recursos del entorno a través de distintos servicios de transporte es modelada por el aspecto de distribución (clase *DistributionAspect* en la Fig.2). La separación de este aspecto permite hacer independiente la participación del agente en una interacción del servicio de transporte utilizado para intercambiar mensajes con otros agentes. Asimismo este aspecto permite que el agente pueda comunicarse con otros agentes a través de distintas plataformas de agentes y servicios de transporte haciéndolo más versátil y adaptable. El aspecto de distribución ha sido implementado siguiendo el patrón de diseño *Facade*[12] proporcionando así un único punto de distribución de mensajes. Esta única instancia, mediante el uso de adaptadores, se encargará de gestionar la distribución de mensajes a través de diferentes plataformas de agentes. Dentro del aspecto de distribución, el envío y recepción de mensajes se realiza a través de un objeto adaptador. Un adaptador debe realizar la interfaz *MTSAdapter* (mostrada en la Fig.2) y se encarga de encapsular aquellas características dependientes de la plataforma en cuestión, como el formato de la envoltura y el protocolo de transporte de mensajes utilizado por la plataforma y la interfaz requerida para acceder a los servicios de la plataforma de agentes. Sobre esta implementación, cuando se configura un agente el programador debe indicar qué adaptadores desea incluir como parte del aspecto de distribución. La implementación de los adaptadores sólo debe realizarse una vez. Aunque aparece sombreado en gris, una vez que disponemos de un adaptador que facilite el acceso a una plataforma no será necesario implementarlo de nuevo. Por tanto, en la mayoría de los agentes se recurrirá a reutilizar las implementaciones disponibles de estos aspectos.

Para completar la descripción de la arquitectura de un agente Malaca tan sólo nos queda describir la función

que desempeña el elemento *Mediador* en la arquitectura del agente. Esta entidad, representada por la clase *Mediator* (en la Fig.2) se encarga, de forma transparente al programador, de crear componentes y aspectos, y de componerlos o "entretejerlos", en terminología de DSOA, en tiempo de ejecución. El *Mediador* desarrolla estas funciones utilizando la información proporcionada por el programador en diferentes documentos XML acerca de la descripción de la arquitectura y configuración del agente. La creación de componentes y aspectos dependerá de la información proporcionada para configurar la arquitectura del agente. Por ejemplo, los conectores no se crean al iniciar el agente, sino que se crean *"bajo demanda"* cada vez que el agente inicia una nueva conversación. Sin embargo, otros aspectos del agente, como el de distribución y representación, se crean y configuran al iniciar el agente. Como parte de la composición de componentes y aspectos el *Mediador* compone los conectores de conversaciones con los componentes funcionales. Durante la ejecución del protocolo de negociación de agentes, el *Mediador* establece la correspondencia entre el servicio solicitado por el conector y los servicios ofrecidos por los distintos componentes incluidos en la arquitectura.

3.2 Implementación de un Agente Vendedor de Libros en Malaca

En esta sección describiremos el proceso de desarrollo del agente vendedor de libros equivalente al agente implementado sobre JADE mostrado en la sección anterior. Partiendo de la especificación dada hemos de localizar en un servidor de componentes COTS, los componentes software que ofrezcan la funcionalidad necesaria para atender peticiones de consulta y de venta de libros. En el peor caso el programador deberá implementar un componente con dicha funcionalidad (clase *BookSellerComponent* en la Fig.2). En concreto este componente está implementado en Java y ofrece esta funcionalidad a través de tres servicios (que se corresponden con métodos de la clase): *OfferRequestsService()*, *PurchaseOrderService()* y *UpdateCatalogue()*. Vemos que, a diferencia de JADE, un único componente encapsula todo el comportamiento relativo al acceso al catálogo de libros. Para que el componente pueda ser incluido y utilizado de forma adecuada en la arquitectura, debemos definir los tres servicios que conforman la interfaz del componente en OWL-S. Parte de esta interfaz, en concreto la descripción del servicio *OfferRequestsService*, se muestra en la Fig.3.

```
<ComponentInterfaceprofile:Service rdf:ID="OfferRequestsService">
  <owls:serviceName> OfferRequestsService </owls:serviceName>
  <owls:hasInput>
    <owls:Input rdf:ID="targetBookTitle">
      <owls:parameterType >&amp;xsd;#string</owls:parameterType>
    </owls:Input>
  <owls:hasOutput>
    <owls:Output rdf:ID="price">
     <owls:parameterType >&amp;xsd;#decimal</owls:parameterType>
   </owls:Output>
  </owls:hasOutput>
</ComponentInterfaceprofile:Service>
```

Figura 3. Descripción en OWL-S del servicio *OfferRequestsService* del componente *BookSellerComponent*.

```
<?xml version="1.0" encoding="UTF-8"?>
<ProtocolDescription …>
  <InterchangedMessages>
    <MessageDescription ID="CFP">
      <Performative>CFP</Performative>
      <ReplyTo>String</ReplyTo>
      <Content contentType="targetBookTitle"
        contentEncoding="STRING"/>
      <Language>STRING</Language>
      <Encoding>String</Encoding>
      <Ontology>bookTrading</Ontology>
      <Protocol>CFP-Protocol</Protocol>
   </MessageDescription>
…
  </InterchangedMessages>

<StateTransitionRule>
  <CurrentState>Idle</CurrentState>
  <Input type="message">cfp</Input>
  <executeTransition>OfferRequest</executeTransition>
  <NextState>End</NextState>
</StateTransitionRule>
```

Figura 4. Descripción de un mensaje y una regla de transición de estados del protocolo *OfferRequestProtocol*.

De la misma manera es necesario describir en XML los dos protocolos de comunicación que debe soportar el agente, ya que su implementación no está entremezclada con el código funcional del agente como en el caso de JADE. El protocolo *OfferRequestProtocol* describe la interacción en la que el agente comprador solicita al agente vendedor el precio de un libro. El segundo protocolo se utiliza para regular la interacción iniciada por el agente comprador para realizar la compra de un libro. La descripción del protocolo en XML incluye la descripción del comportamiento de cualquier participante en la interacción (en este caso correspondería al comprador y a los vendedores). Por motivos de espacio mostraremos sólo la descripción correspondiente al agente vendedor en el primer protocolo citado. Las Figs. 4 y 5 se corresponden con fragmentos de la descripción de este protocolo en XML. Inicialmente, se describen los mensajes intercambiados por los agentes participantes en una interacción. La Fig. 4 muestra la descripción correspondiente al mensaje *cfp* enviado por el agente comprador a los agentes vendedores al iniciar la interacción.

El protocolo de interacción para cada tipo de participante se describe mediante una máquina de estados finita. Esta representación facilita la especificación del protocolo en algún lenguaje de descripción y especificación de protocolos, como por ejemplo SDL o MSC, que permita su análisis mediante alguna herramienta. En este caso la Fig. 4 también muestra cómo se describe el comportamiento del agente con el rol de vendedor mediante una simple regla de transición. Uno de los elementos de la regla identifica la transición que se ejecuta cuando esta regla se activa. Esto ocurre cuando el estado de la conversación es *Idle* y se recibe el mensaje *cfp*. La transición etiquetada como *OfferRequest* describe el comportamiento que lleva a cabo el agente como resultado de la ejecución de dicha transición.

La Fig.5 muestra parte de la descripción de esta transición que determina el comportamiento del agente vendedor cuando recibe un mensaje del tipo *cfp* solicitando el precio de un libro cuyo título se incluye en el contenido del mensaje. El agente consultará el catálogo accediendo al servicio *OfferRequestsService* descrito anteriormente. Si dicho libro se encuentra en el catálogo, enviará un

mensaje *propose* incluyendo el precio como contenido del mensaje. De no encontrarse el libro en el catálogo el agente responderá con un mensaje *refuse* indicando en el contenido que ese libro no está disponible. Hemos de hacer notar que este comportamiento es equivalente al del agente JADE mostrado en la sección anterior. Cómo parte de la transición se hace referencia al servicio *OfferRequestsService,* provocando que se invoque en tiempo de ejecución.

```
<TransitionDescription ID="OfferRequest">
 <CompositeProcess>
  <composedOf>
   <Sequence>
    <components>
      <AtomicProcess ID="OfferRequestsService">
       <hasInput input="targetBookTitle"/>
       <hasOutput output="price"/>
      </AtomicProcess>
      <If-Then-Else>
       <ifCondition>
         <IsNull resource="price"/>
       </ifCondition>
       <then>
        <ReplyMessage resource="refuse">
          <hasContentInput input="not-available"/>
        </ ReplyMessage >
       </then>
       <else>
         <ReplyMessage resource="propose">
           <hasContentInput input="price"/>
         </ ReplyMessage >
       </else>
    </components>
   </Sequence>
  </composedOf>
```

Figura 5. Descripción del comportamiento del agente como parte de una transición.

La descripción de este protocolo de interacción se realizará solamente una vez, y se incluirá en la descripción de todos los agentes involucrados en la interacción. La descripción XML de los protocolos se depositará en una localización accesible para el agente e identificada mediante una URI. Una descripción más completa sobre el esquema XML utilizado en esta especificación de protocolo se encuentra en [13].

A continuación, y como parte del desarrollo del agente es necesario localizar o implementar el aspecto de representación encargado de codificar los mensajes como cadenas (*String*) (formato utilizado por defecto en JADE) y un adaptador que facilite el acceso a una plataforma de agentes JADE para distribuir los mensajes. El adaptador deberá realizar la interfaz Java *MTSAdapter* y proporcionar acceso a una plataforma de agente JADE. Actualmente se proporcionan adaptadores para los servicios de transporte de las plataformas de agentes FIPA-OS, JADE y Zeus. Finalmente, es necesario editar un documento XML que describa los componentes y aspectos incluidos en la arquitectura y su configuración. Para el ejemplo que nos ocupa, y tal y como muestra la Fig. 6, esta información hará referencia al componente *BookSellerComponent*, y a los aspectos de representación y distribución antes mencionados. En la descripción correspondiente al aspecto de coordinación (bajo la etiqueta *Coordination*) se incluyen las referencias a las descripciones de los protocolos de interacción soportados por el agente. La descripción de la comunicación del agente concluye con la descripción de las reglas de aplicación de aspectos que determinan en que orden y cuando se aplican los aspectos incluidos en la arquitectura. La Fig. 6 muestra una regla que define la aplicación de los aspectos de representación y distribución a los mensajes enviados por el agente. La descripción de un agente finaliza detallando el contexto inicial del agente, integrado por las acciones, conversaciones y recursos iniciales del agente al ser creado.

```
<?xml version="1.0" encoding="UTF-8"?>
<AgentDescription>
  <Functionality>
    <ComponentDescription>
      <OntologyID>BookSeller</OntologyID>
      <InterfaceDescription href="BookSellerInterface.xml"
       notation="OWL-S"/>
      <DeploymentInfo href="BookSellerDeployment.xml"
       notation="JavaBean"/>
      <Implementation>BookSellerComponent.class</Implementation>
    </ComponentDescription>
  </Functionality>
  <Communication>
    <Distribution>
      <RoleIdentifier name="Distribucion"/>
      <AspectScope scope="AGENT_SCOPE"/>
      <InstanceIdentifier name="JADE"/>
      <Implementation uri="JadeAdapter.class"/>
    </Distribution>
    <Coordination>
      <RoleIdentifier name="Coordination"/>
      <AspectScope scope="CONVERSATION_SCOPE"/>
      <Implementation uri="ProtocolConnector.class"/>
      <ProtocolDescription href="OfferRequestProtocol.xml"
        notation="ProtocolSchema.xsd"/>
    </Coordination>
…
    <ACLRepresentation>
      <RoleIdentifier name="Representacion"/>
      <AspectScope scope="AGENT_SCOPE"/>
      <InstanceIdentifier name="acl.rep.std.string"/>
      <Implementation uri="StringRepresentationAspect.class"/>
    </ACLRepresentation>
<ApplicationRule>
  <InterceptionPoint>SEND_MSG</InterceptionPoint>
  <ApplyAspect Role="Representation" RoleInstance="acl.rep.std.string"/>
  <ApplyAspect Role="Distribution" RoleInstance="JADE"/>
</ApplicationRule>
…
  <InitialContext>
…
  </InitialContext>
</AgentDescription>
```

Figura 6. Descripción en XML del agente Vendedor de Libros desarrollado sobre la arquitectura Malaca.

4 Comparación de ambas implementaciones.

A la vista de los procesos de desarrollo comentados en las secciones previas un punto diferenciador relevante es que Malaca no impone al desarrollador la implementación de ningún elemento del marco de trabajo. Considerando el caso mejor ni siquiera será necesario desarrollar los componentes proveedores de la funcionalidad del agente, ya que si se desarrollan diferentes sistemas de agentes dentro de un mismo dominio de aplicación es posible que hayan sido desarrollados. Tampoco supone un gran esfuerzo de programación incorporar al agente un nuevo protocolo de comunicación. Esta tarea requiere tan sólo describir en XML dicho protocolo. Supongamos que el protocolo de interacción *OfferRequestProtocol* necesita ser modificado para utilizar el protocolo *FIPA-Request*. Como ya comentamos, en JADE los protocolos para la venta de libros se encuentran "entremezclados" con la funcionalidad específica de acceso al catálogo de libros, por lo tanto modificar el protocolo de interacción implica buscar y modificar el código en la clase correspondiente. En cambio, para el agente Malaca tan sólo será necesario modificar la descripción XML del protocolo.

Supongamos ahora que el agente, en lugar de acceder a un catálogo local representado por una tabla *hash* (como es el caso), tuviera que acceder a un catálogo *on-line* almacenado en una base de datos. Para el agente implementado en JADE la modificación afectaría a los tres comportamientos implementados, es decir, sería necesario modificar las tres clases, recompilar y comprobar que no se ha producido

ningún error ni inconsistencia durante la actualización. Sin embargo, para el agente implementado con Malaca este cambio sólo supone sustituir en su descripción en XML la implementación del componente por otra que se adapte a la nueva funcionalidad, es decir que modele el catálogo mediante un acceso a una base de datos. Como podemos observar, el impacto de la actualización es menor.

El cambio en la especificación puede afectar no sólo a la funcionalidad, sino también a las propiedades extra-funcionales del agente. Supongamos ahora que se desea registrar mediante una traza el envío y recepción de mensajes en un fichero de *log*. En el caso de la implementación en JADE esto supone, de nuevo, inspeccionar el código y añadir el código correspondiente a la escritura en fichero cada vez que se realiza una llamada a los métodos *send()* y *receive()*. Una tarea, por otra parte, bastante tediosa y repetitiva. En el caso del agente Malaca este cambio en la especificación se traduce en la identificación e incorporación de un nuevo aspecto en la arquitectura, el aspecto de traza. Cómo ocurre con el aspecto de representación, si no dispusiéramos de una implementación adecuada del aspecto necesitaríamos implementar su comportamiento en una clase que extienda a la clase *Aspect* definida por el marco de trabajo Malaca. Una vez implementado o localizado el aspecto de traza, tan sólo hay que incluir su descripción y configuración como parte de la descripción del agente, en incorporar en la reglas de composición de aspectos su aplicación cuando al agente envía y recibe un mensaje. Por otra parte, hemos de puntualizar que la configuración de un agente Malaca no tiene que permanecer fija una vez que el agente es creado. Es posible registrar nuevos componentes o actualizar los registrados en tiempo de ejecución para incorporar nuevas versiones, o puede "aprender" nuevos protocolos de negociación cargando nuevas descripciones [13], o incorporar nuevos adaptadores, codificadores y decodificadores sin necesidad de suspender la ejecución del agente o reemplazarlo por otro que soporte nueva funcionalidad.

5 Conclusiones

En este trabajo presentamos una arquitectura de agentes software basada en componentes y aspectos que proporciona una infraestructura para construir agentes a partir de componentes software reutilizables. Esta arquitectura combina técnicas propias del DSBC y DSOA y separa en entidades distintas la funcionalidad, la coordinación del agente, la representación y la distribución de mensajes. De esta forma se consiguen agentes software más flexibles, adaptables y reutilizables, características que ayudan a mejorar sensiblemente su proceso de desarrollo. El proceso de desarrollo de un nuevo agente se reduce, en el mejor de los casos, a la descripción en un documento XML de la composición inicial del agente en cuanto a funcionalidad, protocolos de coordinación y propiedades que especifican la comunicación con otros agentes. La utilización de XML para representar la descripción y configuración de los elementos arquitectónicos de un agente facilita y simplifica el desarrollo de agentes software, ya que no exige al programador el conocimiento de un lenguaje de programación y/o de un marco de trabajo concreto.

Referencias

[1] The Zeus Agent Building Toolki ,BtexaCT,. http://193.113.209.147/ projects/agents/zeus

[2] FIPA-OS, Emorphia. http://fipa-os.sourceforge.net/

[3] Java Agent Development Framework, TILAB, http://JADE.cselt.it

[4] G.T. Heineman, W.T. Councill, "Component-Based Software Engineering: Putting the Pieces Together", Addison Wesley,(2001).

[5] C. Szyperski, Component Software: Beyond Object-Oriented Programming, Addison Wesley, (1998).

[6] M. Amor, L. Fuentes, J.M. Troya, "Integración de Servicios Web mediante un Modelo Composicional de Agentes Software", JITEL 2003.Gran Canaria, 15-17 de Septiembre (2003).

[7] Kickzales G., Lamping J., Mendhekar J., Maeda C., Videira Lopes C, Loingtier J., Irwin J. Aspect-Oriented Programming. Proceedings of the European Conference on Object- Oriented Programming (ECOOP´97), Finlad. Springer-Verlag LNCS 1241, Junio (1997).

[8] Aspect-Oriented Software Development. http://www.aosd.net

[9] FIPA. Foundation for Intelligent Physical Agents. http://www.fipa.org.

[10] OWL-S, OWL-based Web Service Ontology. http://www.daml.org/services/owl-s/1.0/

[11] M. Amor, L. Fuentes, L. Mandow y J.M. Troya, "Building Software Agents from Software Components" Lecture Notes in Artificial Intelligence, pp. 222-231, vol. 3040. (2004).

[12] Gamma, E. et al. Design Patterns Elements of Reusable Object-Oriented Software, Addison Wesley editors (1995).

[13] M. Amor, L. Fuentes, J.M. Troya, "Training Compositional Agents in Negotiation Protocols" in Integrated Computer-Aided Engineering International Journal. Vol. 11, No. 2 pp. 179-194 (2004).

Uso de Source-Specific Multicast en aplicaciones multimedia interactivas multipunto[1]

Vicente Sirvent, Gabriel Huecas, Carlos Barcenilla, David Fernández, Juan Quemada
Departamento de Ingeniería Telemática, Universidad Politécnica de Madrid
E.T.S.I. de Telecomunicación, c/ Ciudad Universitaria, s/n, Madrid, 28040, España
E-m@il: {sirvent, ghuecas,barcenilla,dfernandez,jquemada}@dit.upm.es

Abstract. *A key aspect in the transmission of multipoint interactive media is the optimal distribution of the multimedia traffic among the collaborating terminals. The Source-Specific Multicast (SSM) service model allows the distribution of multiple receiver packets through single logical topologies, whilst solving many of the problems caused by classical Any Source Multicast (ASM). In this paper, we describe the use of SSM for the optimal distribution of multimedia flows in an interactive multipoint service. The solution to the source address discovery inherent to SSM, based on a source address distribution protocol, is also described. This protocol is commonly used in sessions with one or few known sources, meanwhile having a large number of dynamic, a priori unknown, media sources. We also present a prototype of a real-time multipoint audio/video/data conferencing application, and describe the transition scenarios used to validate it over both IPv4 and IPv6.*

1. Introducción[1]

La transmisión de medios interactivos multipunto se ha convertido en la actualidad en un campo con gran actividad investigadora debido a la gran variedad de usos que ofrece, al aumento constante del segmento de población que hace uso de esta tecnología, y a la constante mejora de las actuales redes domésticas. Entre los servicios más importantes basados en la transmisión de medios interactivos encontramos las aplicaciones colaborativas interactivas que requieres protocolos eficientes para las comunicaciones n-a-n, donde varios flujos multimedia se han de enviar de cada miembro de un grupo al resto, con latencia y jitter (variación en el retardo) mínimos y un uso óptimo del ancho de banda de la red.

El uso del modelo tradicional de multicast ASM (Any Source Multicast) resulta especialmente útil en la distribución de tráfico multimedia interactivo entre múltiples terminales [14][15][10]. Sin embargo, ASM adolece de ciertos inconvenientes que han impedido su empleo generalizado [1].

El modelo de servicio SSM (Source-Specific Multicast), simplifica el servicio de distribución restringiéndose a una fuente específica por grupo multicast. Sin embargo, SSM ha sido diseñado principalmente para ser utilizado en aplicaciones con una o muy pocas fuentes conocidas, en una distribución 1-a-n. Por ello se requiere una capa de adaptación que gestione las direcciones de las fuentes y grupos necesarios en las comunicaciones n-a-n..

En este artículo se describe la implementación realizada sobre el sistema multimedia interactivo multipunto, ISABEL. También se describe un protocolo de distribución de direcciones fuentes que permite el uso de SSM en aplicaciones con múltiples fuentes dinámicas y su implementación en ISABEL. Posteriormente se describen la serie de pruebas sobre escenarios reales tanto en IPv4 como en IPv6, orientadas a validar el uso de SSM, el algoritmo de distribución de fuentes y la implementación en ISABEL. Finalmente, se incluye el estudio realizado y su evaluación del tráfico generado y del retraso introducido por la red en las distintas pruebas.

[1] Este trabajo ha sido parcialmente subvencionado por el proyecto Euro6IX (European IPv6 Internet Exchanges Backbone), IST-2001-32161.

2. SSM: Antecedentes y trabajos relacionados

2.1. Source-Specific Multicast

Source-Specific Multicast está ampliamente basado en el protocolo Express de Holbrook y Cheriton [6][7]. SSM[1][2] es un modelo de servicio desarrollado mediante el uso de los protocolos PIM-SM[3] y IGMPv3/MDLv2 [4][5] que, al igual que el protocolo de Holbrook y Cheriton, ofrece un servicio en la capa de red que permite a un terminal recibir flujos desde una dirección multicast (G) con origen en una dirección IP fuente (S), mediante la suscripción a "*canales*" identificados por el par (S, G), permitiendo de esta manera seleccionar el origen de los datos recibidos por los terminales.

IGMPv3 para IPv4 y MDLv2 para IPv6 son los protocolos que dan soporte a la suscripción/baja de canales. El árbol de envío de los datagramas IP se crea con raíz en la fuente (S), y se va construyendo usando el protocolo PIM-SM (Sparse Mode). El modelo de servicio SSM resuelve algunos de los inconvenientes que presenta el uso de ASM[1]:

a) Localización de direcciones: SSM define *canales* basándose en fuentes, por ejemplo el canal (S_1, G) es distinto del canal (S_2, G), donde S_1 y S_2 son direcciones de fuente distintas, y G es una dirección de grupo SSM. Este hecho resuelve el problema de utilizar direcciones SSM globales y la necesidad de coordinar la asignación de direcciones

entre las distintas fuentes, reduciendo al ámbito de cada fuente la responsabilidad de solventar las colisiones de direcciones producidas en la creación de sus canales.

b) Control de acceso: SSM resuelve parcialmente el problema del control de acceso que conlleva el uso de ASM, tanto en destino como en origen. En cuanto al destino, cuando un receptor se suscribe a un canal (S, G), únicamente recibe datos generados por la fuente S. Y respecto al origen, cuando una fuente transmite a un canal (S, G) se asegura de que ninguna otra fuente emitirá por el mismo canal. Esto dificulta en gran medida la tarea de suplantar una fuente.

c) Gestión de Fuentes conocidas: SSM únicamente requiere el uso de árboles con raíz en la fuente; esto elimina la necesidad del uso de una infraestructura de árbol compartido. Esto implica que ni el árbol basado en RP de PIM-SM ni el protocolo MSDP son requeridos. Por tanto, la complejidad de la estructura de enrutado de SSM es baja, lo que facilita el despliegue de este modelo de servicio.

En cuanto a la compatibilidad de uso de SSM junto a ASM, IANA ha reservado para el uso de SSM el rango de direcciones 232/8 en IPv4, mientras que IPv6 contempla en el protocolo un rango reservado para direcciones de grupo (FF3x::/96). Se pretende que exista cierta interoperabilidad entre SSM y ASM, para ello, SSM únicamente podrá trabajar en el rango de direcciones reservadas para dicho protocolo y ASM se ofrecerá en el rango multicast restante donde las suscripciones de canales serán del estilo (*, G), es decir, en ASM los receptores se suscriben a todos los "*canales*" de una dirección de grupo. En cualquier caso, un receptor podrá suscribirse a canales (S, G) con direcciones de grupo fuera del rango SSM, aunque no se garantizará que el comportamiento cumpla con el modelo SSM.

2.2. Multicast con múltiples emisores

Para sesiones con múltiples emisores Holbrook y Cheriton proponen dos soluciones. En la primera, se propone el uso de un único canal para toda la sesión, donde todos los emisores transmitan a través de dicho canal, esta solución tiene la ventaja de que los receptores únicamente deben suscribirse a un único canal para recibir múltiples flujos, pero las desventajas del aumento del retraso entre un emisor y un miembro del grupo al no hacer uso del camino más corto.

La segunda solución propone el uso de un canal diferente para cada emisor, esto mejora el retraso entre una fuente y un miembro del grupo, aunque aumenta la complejidad de enrutado necesario por la aplicación, además de que surge la necesidad de distribuir los distintos canales disponibles entre los miembros del grupo.

3. Uso de SSM para transmisión de flujos multimedia interactivos multipunto

Para el uso del modelo de servicio SSM en transmisiones de flujos multipunto proponemos dos escenarios topológicos de red:

a) Session Relay: En esta topología, mostrada en la Fig. 1, todos los transmisores envían sus flujos a un servidor de medios (Session Relay) (S_1), el cual reenvía dichos flujos a los miembros del grupo SSM (G) a través del canal (S_1, G). De esta manera los receptores únicamente deben suscribirse al canal (S_1, G) para recibir los flujos generados por múltiples fuentes. Esta solución tiene, sin embargo, dos desventajas importantes, ya identificadas en [1]: primero, no se hace un uso óptimo del ancho de banda de la red al producirse replicas de paquetes en el servidor de medios; segundo, hay un aumento del retraso entre un emisor y un miembro del grupo al no usarse el camino más corto.

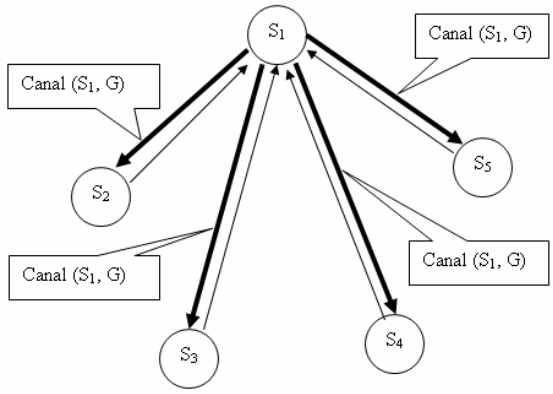


Figura 1: Topología de Servidor de Medios (Session Relay).

b) Nube multicast SSM: En esta topología, mostrada en la Fig. 2, cada una de las fuentes (S) hace uso de un canal (S, G) diferente. Por lo tanto, cada uno de los receptores debe suscribirse a cada uno de los distintos canales utilizados. Esta topología tiene toda la funcionalidad de la topología de nube multicast ASM y aporta todas las ventajas que posee SSM sobre ASM. Sin embargo, es una solución más compleja, pues cada receptor necesita suscribirse a cada uno de los emisores específicamente, por lo que debe conocer las direcciones de cada uno de los emisores.

Esta última solución es la seleccionada en nuestra propuesta, por lo que junto a esta topología debemos hacer uso de algún tipo de algoritmo que posibilite a los distintos miembros de la dirección de grupo SSM (G) conocer las distintas direcciones de las múltiples fuentes emisoras (S_1,…, S_n).

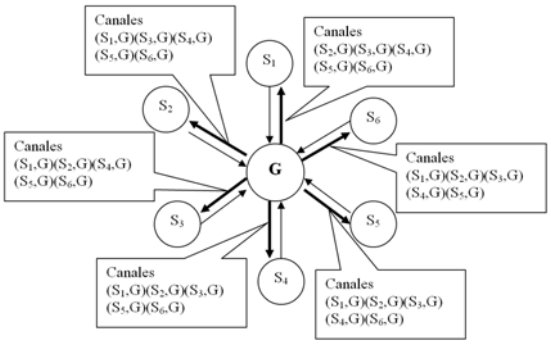


Figura 2: Topología de nube multicast SSM.

Esta última solución es la utilizada en nuestra propuesta, por lo que junto a esta topología debemos hacer uso de algún tipo de algoritmo que posibilite a los distintos miembros de la dirección de grupo SSM (G) conocer las distintas direcciones de las múltiples fuentes emisoras.

3.1 Descubrimiento de fuentes: protocolo de Zona SSM

Para poder hacer uso de la *nube multicast SSM* es necesario que todos los receptores puedan suscribirse y darse de baja de cada uno de los canales (S, G) utilizados por las distintas fuentes. Con este fin, se propone el uso del siguiente algoritmo de descubrimiento de fuentes (S):

a) Definiremos *zona SSM* como el conjunto de *canales* (S, G) que posibilitan el intercambio de tráfico entre distintos terminales. Cada *zona SSM* posee un único *padre* y una única dirección de grupo SSM (G) que la diferenciarán del resto de *zonas SSM*.
b) Supondremos un grupo de sistemas finales pertenecientes a una misma sesión, donde cada uno de ellos desea transmitir al resto de terminales. Un grupo de terminales pertenecientes a una misma sesión hará uso de una misma zona SSM. Las direcciones de los terminales son (S_1,…, S_n).
c) Uno de los terminales actuará como padre de la zona SSM, el resto de los participantes, como hijos. Todos los hijos han de conocer la dirección IP del padre.
d) Cada cierto tiempo, t_1, si un terminal hijo desea hacer uso de la zona SSM, mandará un datagrama IP al padre de la zona SSM, indicando su dirección IP.
e) El padre de la zona SSM, almacenará las direcciones IP recibidas durante cierta cantidad de tiempo t_2.
f) Posteriormente, el padre de la zona SSM actualizará su lista de hijos de la zona SSM y enviará a cada dirección IP de la lista un datagrama IP conteniendo dicha lista de direcciones además de la suya en primera posición.
g) Cada hijo de la zona SSM recibirá de esta manera una lista actualizada de todas las posibles fuentes (S) de la *zona SSM*, posibilitando que cada intervalo de tiempo todos los *hijos* sean capaces de actualizar sus listas de canales (S, G) mediante suscripción o baja de los mismos.
h) En caso de caída del *padre* de una *zona SSM*, el *hijo* cuya dirección IP se encuentre en el primer lugar de la lista de direcciones tomará el papel de *padre* y emitirá una nueva lista de direcciones fuentes (S) colocando su propia dirección en la cabeza de la lista, de esta manera el resto de *hijos* de la *zona SSM* detectarán el cambio de *padre* y el algoritmo podrá seguir funcionando de manera normal.
i) Si un padre caído regresa a la zona SSM, tomará el rol de hijo, ya que normalmente habrá sido relevado de su puesto por un hijo.

Entre las características de este algoritmo destacamos:

- La distribución de las direcciones no requiere una transmisión fiable. Es suficiente informar de las direcciones IP por UDP.
- Ofrece la posibilidad de que coexistan distintas aplicaciones que hagan uso de manera simultanea una única dirección de grupo (G), ya que en cada aplicación y cada sesión puede hacer uso de un zona SSM distinta.
- En caso de caída del padre, el algoritmo busca un nuevo padre con el fin de evitar la caída de la sesión.
- El tiempo de actualización de las fuentes t_2 es variable, y debe ser elegido de tal manera que tanto el ancho de banda consumido por los paquetes debidos al algoritmo de distribución, que generalmente es muy pequeño, como el tiempo de actualización suscripción/baja de canales (S, G) sea aceptable. Para el estudio de este tiempo de actualización se implementó dicho algoritmo en la aplicación ISABEL y se estudiaron los resultados obtenidos.

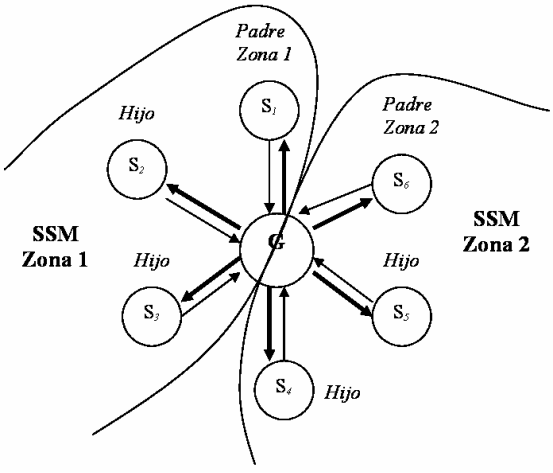


Figura 3: Varias sesiones utilizando nube multicast SSM y zonas SSM distintas.

La Fig. 3 muestra como diferentes terminales pueden pertenecer a diferentes zonas SSM, incluso simultáneamente, ya que los diferentes grupos de SSM usan la dirección de grupo G junto con la propia dirección IP (S) del emisor.

4. Adaptación de la aplicación ISABEL para el uso sobre SSM

Con el fin de poder realizar pruebas sobre escenarios reales, se ha implementado la propuesta de uso de nube multicast SSM junto al protocolo de *zona SSM* en la aplicación ISABEL [8][9]. En esta aplicación, orientada a tele-educación, tele-trabajo y tele-conferencia, todos los nodos colaborando en una sesión pueden recibir simultáneamente los flujos multimedia del resto de los participantes, por lo que la óptima distribución de flujo es un requisito imprescindible [10][11]. A continuación se describe el procedimiento que se ha seguido para implementar de manera exitosa la propuesta.

La implementación ha sido realizada en la aplicación ISABEL v4.8, en lenguaje C++, sobre el sistema operativo Linux Suse 9.1 [12] cuyo núcleo soporta los protocolos IGMPv3 y MDLv2 necesarios para la suscripción/baja de *canales.*

Se trabajó paralelamente en dos líneas de desarrollo. La primera de ellas encargada de implementar las clases que encapsularan los sockets UDP y los métodos de suscripción/baja de *canales* (S, G) que ofrece el servicio SSM, la segunda de ellas encargada de implementar el protocolo de *zona SSM.*

4.1. Detalles de implementación

El modelo de servicio SSM ofrece dos APIs diferentes orientadas a facilitar la programación de diferentes tipos de aplicaciones. Ambas permiten trabajar tanto sobre IPv4 como IPv6.

a) Modo básico (delta-based) API: Proporciona un API básico sobre las opciones de los sockets para la suscripción a canales (S, G) tanto sobre IPv4 como IPv6 , pero no permite la suscripción a varios canales (S, G) de manera simultanea.
   Entre las aplicaciones típicas en las que se usa el modo básico SSM, hay que destacar la radio y la televisión a través de Internet, donde todas las emisoras emiten a una misma dirección de grupo (G), y cada receptor se suscribe a un único canal (S, G) dependiendo del canal que quieran recibir en cada momento.
b) Modo completo (full-state) API: Trabaja directamente sobre los filtros de fuentes del socket, sobre IPv4 e IPv6, añadiendo o quitando las fuentes (S) a las que queramos suscribirnos o darnos de baja. Este API permite que un destino pueda suscribirse a distintas fuentes de manera simultánea.
   Como aplicaciones donde es necesario el uso del modo completo SSM tenemos todas aquellas en las que un receptor necesite suscribirse de manera simultanea a múltiples canales (S, G), entre las que se encuentra la aplicación en la cual se ha realizado nuestra implementación de SSM.

Destaca el hecho de ciertas incompatibilidades entre algunas cabeceras del núcleo 2.6.5 que obligaron a definir las estructuras necesarias para el correcto uso de SSM. En posteriores versiones del núcleo (2.6.8 en adelante) los problemas son subsanados, por lo que no es necesaria la definición de dichas estructuras de manera explícita.

Respecto del protocolo de zona SSM, en el apéndice I se describe el algoritmo mediante seudo código.

5. Escenarios de prueba y resultados

Tras realizar la implementación del modelo de servicio SSM en la aplicación ISABEL, fueron diseñados un conjunto de escenarios de pruebas con el fin de validar el trabajo realizado sobre la aplicación. En esta sección se realiza una descripción detallada de dichos escenarios. Esto se ha llevado a cabo con tres escenarios.

El primer escenario de prueba se muestra en la Fig. 4, en el que diversas máquinas, con direcciones IPv4 e IPv6, participan en un evento con la aplicación ISABEL. Este escenario se utilizó para validar el protocolo de distribución de fuentes, probando tanto en IPv4 como en IPv6. Se comprobó el correcto uso de los servicios ofrecidos por IGMPv3 en IPv4 y MDLv2 en IPv6 para la suscripción/baja de los participantes a los diferentes *canales* (S, G) utilizados en la sesión de ISABEL.

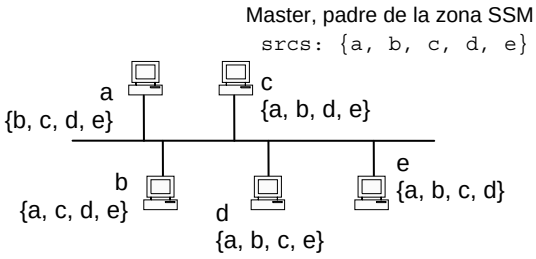


Figura 4: Primer escenario: validación del protocolo de zona SSM (IPv4 e IPv6)

También hemos verificado y puesto a prueba el correcto funcionamiento del protocolo de zona SSM mediante diferentes tests donde los participantes se unían y dejaban la sesión. Hemos comprobado como un tiempo de actualización de las direcciones de fuentes en los receptores de unos pocos segundos resulta una selección correcta, de tal manera que el protocolo de zona SSM funciona correctamente permitiendo a los distintos participantes unirse y

dejar sesiones rápidamente a la vez que su uso supone una sobrecarga mínima en la red.

La caída del padre de la zona SSM y su sustitución por el primer hijo de la lista también fue probado con éxito, esto supone una importante ventaja: en caso de caída del nodo padre, no se verá comprometida la sesión multimedia, es más, dicho nodo podrá posteriormente volver a unirse a la sesión.

El segundo escenario consiste en tres terminales, (S_1, S_2, S_3) participando en una sesión de ISABEL. Cada terminal se encuentra localizado diferentes redes Ethernet (lan_1, lan_2, lan_3) conectados mediante tres routers Cisco 7204 (r_1, r_2, r_3). Los routers se encuentran unidos entre sí mediante enlaces ATM punto-a-punto. La Fig. 5 muestra el esquema de red utilizado en este escenario.

La batería de pruebas en sesiones de intercambio de flujos multimedia multipunto entre tres participantes diferentes *(S_1, S_2, S_3)*: uno de los nodos (S_1) actúa como padre P de la zona SSM, mientras los otros dos (S_2, S_3) lo hacen como hijos, formando así una zona SSM (S_1, G). Tanto IPv4 como IPv6 fueron utilizados sobre este escenario.

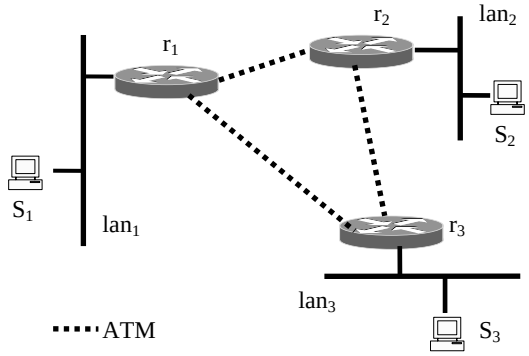


Figura 5: Segundo escenario: Validación a través de diferentes redes.

Básicamente, el fin de este segundo escenario fue el de verificar el correcto funcionamiento de la aplicaciones sobre una pequeña red formada por tres routers, así como el correcto funcionamiento del modelo de servicio SSM en los routers *Cisco 7200*.

Finalmente, diseñamos un tercer escenario de pruebas con fin de probar la aplicación utilizando SSM sobre un escenario real de red. En este escenario hemos utilizado el servicio piloto de SSM sobre la red pan-Europea del proyecto Euro6IX [13].

En este escenario, distintos países europeos se unieron a una misma sesión de ISABEL a través de la red global. La Fig. 6 muestra un mapa de la sesión, donde se diferencian los terminales conectados mediante unicast (directamente) de los conectados mediante SSM.

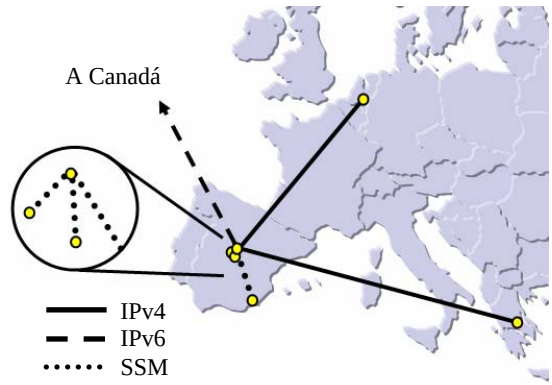


Figura 6: Conexiones en evento real de validación de SSM.

Con el propósito de validar nuestra propuesta en escenarios de red más complejos el 26 de mayo del 2005, pudimos comprobar el funcionamiento de ISABEL haciendo uso de SSM mediante la organización de un evento público sobre la topología de red mostrada en la Fig. 6. La sesión estuvo formada por 7 participantes, 4 de ellos haciendo uso de SSM: Universidad Politécnica de Madrid (UPM), Universidad de Murcia (UMU), Telefónica I+D (TID) y Consulintel (CON). El resto de los participantes fueron Communications Research Centre Canada (CRC), National Centre of Scientific Research of Greece (DEM) y Université Libre de Bruxelles (ULB), los cuales participaron mediante conexiones directas a una MCU localizada en la Red Española de I+D (redIRIS) .

La Fig. 7 muestra en detalle la topología de red y los routers utilizados en este tercer escenario. La mayoría de las conexiones se realizaron en IPv6 y la parte de SSM se configuró para soportar dicho servicio de forma nativa. En particular, la Universidad de Murcia (UMU) se unió a la zona SSM a través de un túnel IPv6 sobre IPv6.

El escenario de pruebas fue configurado en la parte Española de la red Euro6IX donde UPM, CON y TID funcionaron sobre conexiones multicast IPv6 nativas a través de MAD6IX Exchange. Es importante el hecho de que, mientras que la mayoría de fabricantes de routers y Redes Nacionales de Investigación (NREs) se posicionan claramente a favor del soporte de SSM como parte de grupo de servicios ofrecidos por PIM-SM, en la actualidad estos servicios no se encuentran claramente definidos, y tuvimos que hacer frente en este tercer escenario de pruebas con graves problemas de interoperabilidad entre routers de distintos fabricantes. Una vez resueltos los problemas de interoperabilidad, la suscripción a los distintos canales SSM se pudo realizar y la sesión transcurrió de manera exitosa.

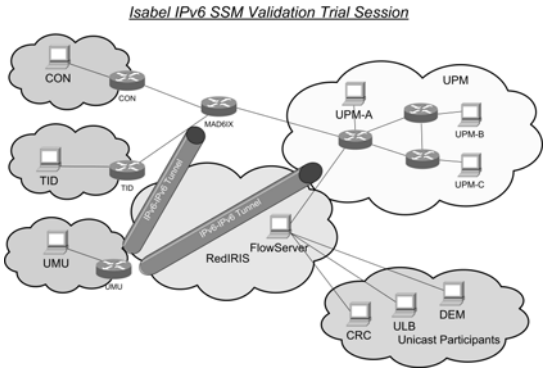


Figura 7: Esquema de red utilizado en el tercer escenario

La Tabla I muestra las medidas que obtuvimos en el tercer escenario en cuanto a la variación de retardo para cada uno de los nodos participantes, tanto los que hacían uso de SSM como para los que realizaban conexiones directas a través de la MCU. Cada participante mandó un flujo de audio y uno de video al resto de terminales además de un flujo de datos para dar soporte a la presentación de diapositivas.

Pudimos observar como la sobrecarga de ancho de banda introducida por la distribución de direcciones fuente supuso únicamente el 0,32% (4800 bps).

Los valores medios de la variación del retardo fueron muy bajos, incluso en grandes distancias como Canadá. Los resultados que obtuvimos de los participantes de la zona SSM fueron comparables con los obtenidos en conexiones directas.

Nodo	Variación en el retardo (ms.)
Participantes de zona SSM	
Consulintel	5.58
Telefónica	6.10
Univ. Murcia	3.05
UPM-A	3.32
UPM-B	4.80
UPM-C	3.40
Participantes Unicast	
CRC (IPv6)	11.30
DEM (IPv4)	1.89
ULB (IPv4)	7.59

Tabla I. Variación en el retardo (jitter) en ms. para participantes de zona SSM y unicast

6. Conclusiones

Este artículo presenta un estudio e implementación de un esquema de transmisión multimedia eficiente a través de múltiples participantes pertenecientes a una misma sesión colaborativa mediante el uso del modelo de servicio SSM. Hemos implementado de manera exitosa dicho esquema en una aplicación real denominada ISABEL.

Como fue anteriormente mencionado, esta aplicación requiere transmisiones multimedia interactivas n-a-n, de tal manera que el servicio básico ofrecido por SSM – orientado a transmisiones 1-a-n – debe ser utilizado conjuntamente con un protocolo de distribución de direcciones fuente denominado algoritmo de la zona SSM.

Por un lado, el uso de SSM resuelve los problemas que presenta el uso de ASM, los cuales han impedido su despliegue a nivel global, mientras que mantiene la simplicidad y eficiencia que ofrece multicast en la transmisión de flujos multimedia.

Con la adaptación de una aplicación real y su posterior uso en diferentes escenarios de pruebas, hemos comprobado como el uso de SSM como medio de transporte para este tipo de transmisiones, requeridas generalmente por aplicaciones en tiempo real colaborativas interactivas, es posible y muy adecuado.

A pesar de que es necesaria la realización de más pruebas en escenarios reales, los resultados obtenidos en cuanto a la variación del retardo muestran que el uso de SSM no produce ningún impacto en este sentido con respecto al uso de conexiones directas. El ancho de banda introducido en la distribución de direcciones fuente resulta insignificante comparado con el ancho de banda total utilizado generalmente en este tipo de aplicaciones.

Tal vez, considerando los resultados obtenidos a lo largo del trabajo realizado y presentados en este artículo, deberíamos plantearnos la posibilidad de generalizar el uso de SSM en aplicaciones multimedia multipunto. Sin embargo, este paso requiere un mayor soporte del modelo de servicio SSM por parte de las redes y proveedores a nivel global.

Apéndice I
Detalles de implementación del protocolo de Zona SSM

En el protocolo de *zona SSM* para la distribución de fuentes se distinguen dos comportamientos distintos dependiendo del rol que desempeñe cada terminal en una *zona SSM*. El siguiente texto muestra en lenguaje pseudo-código el protocolo de distribución de fuentes para un terminal *padre* y uno *hijo* de una *zona SSM*:

```
process PADRE
lista.añadir(miDir)
call DISTRIBUIR_FUENTES(lista)
while (1)
  time1 = time2 = HORA_SISTEMA;
  /*recolectar fuentes*/
  while(time2 – time1 < CICLO)
    leer_dir(dir)
    if not(lista_tmp.tiene(dir)) then
    lista_tmp.insertar(dir)
    end if
    time2 = HORA_SISTEMA
  end while

  /*borrar fuentes que no
```

```
    han enviado su direccion */
  for each dir in lista
    if not(lista_tmp.tiene(dir)) then
      lista.borrar(dir)
    end if
  end for

  /*Añadir nuevas fuentes */
  for each dir in lista_tmp
    if not(lista.tiene(dir)) then
      lista.añadir(dir)
    endif
  end for

  /*distribuir fuente
    actualizada a los hijos */
  call DISTRIBUIR_FUENTES(lista)
end while
/*volver a recolectar fuentes */
end process

process HIJO
while(true)
  enviar_a_padre(dir)
  recibir(lista_tmp)
  esperar(timeout)
  /*Si cae el padre y soy
    primogénito paso a ser padre*/
  if (lista_tmp==NULL) then
    if (lista.head()==miDir) then
      call PADRE
    end if
  end if

  /*dar de baja fuentes */
  for each dir in lista
    if not (lista_tmp.tiene(dir)) then
      lista.borrar(dir)
      leaveSSM(dir,grupo)
    end if
  end for

  /*suscribirse a nuevas fuentes */
  for each dir in lista_tmp
    if not(lista.tiene(dir)) then
      lista.añadir(dir)
      joinSSM(dir,grupo)
    end if
  end for
end while
end process
```

7. Referencias

[1] RFC 3569 - An Overview of Source-Specific Multicast (SSM) S. Bhattacharyya, Ed.

[2] RFC 3678 - Socket Interface Extensions for Multicast Source Filters

[3] RFC 2362 - Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification

[4] RFC 1112 - Host extensions for IP multicasting

[5] RFC 2710 - Multicast Listener Discovery (MLD) for IPv6

[6] D.R. Cheriton and H.W. Holbrook, "EXPRESS Multicast: Making Multicast Economically Viable," en *ACM SIGCOMM*, August 1999.

[7] H.W. Holbrook and D.R. Cheriton. **"**IP Multicast Channels: Express Support for Large-scale Single-source Applications" Department of Computer Science, Stanford University, *ACM SIGCOM*M, September 1999.

[8] J. Quemada, T. de Miguel, S. Pavón, J. Salvachúa, M. Petit, T. Robles, G. Huecas, F. Echevarrieta, E. Castro. "La Aplicación ISABEL" en Seminario del Programa Nacional de Aplicaciones y Servicios Telemáticos, ISBN: 84-95075-27-X. 2000

[9] The Isabel Application: www.isabel.dit.upm.es

[10] J. Quemada, T. Miguel, S. Pavón, G. Huecas, T. Robles, J. Salvachúa, M.J. Perea, E. Moro, D.A. Acosta, J.A. Fernández, F. Escribano, A. Diaz, J.L. Fernández, J.A. Sánchez, J. Sedano, M. Gómez. "Isabel: An Application for real time Audience Interconnection over the Internet." Terena Networking Conference 2004 Rhodes, Greece. June 7-10, 2004.

[11] J. Quemada, T. de Miguel, E. Castro, S. Pavón, G. Huecas, T. Robles, J. Salvachúa, E. Apolinario, J. Sedano, M.J. Perea. "Isabel Distribution of the Madrid Global IPv6 Summit 2002 over an IPv6 Transition Network" en 2003 Symposium on Applications and the Internet Workshops, pps. 200-203 ISBN:0-7695-1873-7 IEEE Computer Society. (SAINT 2003 Workshops) Orlando, Florida, U.S.A. 27-31 Enero 2003.

[12] SuSe Linux : www.suse.com

[13] Proyecto Euro6IX (European IPv6 Internet Exchanges Backbone): www.euro6ix.org

[14] I. Matta, M. Eltowaissy, K. Lieberherr "From CSCW Applications to Multicast Routing: An Integrated QoS Architecture Communications". 1998 ICC'98 Conference Record. IEEE International Conference Vol. 2, pps. 880-884, 7-11 June 1998.

[15] J.L. Moreno, A. Azcorra, D. Larrabeiti, T. de Miguel, M. Alvarez-Campana "Multimedia-multiparty service support in ATM wide area networks" IEEE Conference on Protocols for Multimedia Systems- Multimedia Networking (PROMSS MmNet'97) Santiago, Chile, 1997.

Punishing Manipulation Attacks with the Mobile Agent Watermarking Approach

Oscar Esparza Miguel Soriano Jose L. Muñoz

Abstract *Mobile agents are software entities consisting of code, data and state that can migrate autonomously from host to host executing their code. Despite their benefits, security issues restrict the use of code mobility. The protection of mobile agents against the attacks of malicious hosts is considered the most difficult security problem to solve in mobile agent systems.*

This paper introduces some techniques that aim to solve the problem of the malicious hosts. First of all, a lightweight attack detection technique is explained. This technique is based on embedding a watermark into the agent, i.e. Mobile Agent Watermarking. Additionally, this paper introduces the protocols that can be used to punish the malicious host by using a Third Trusted Party, the Host Revocation Authority.

1 Introduction

Mobile agents are software entities that can migrate autonomously from host to host performing actions on behalf of a user. Mobile agents not only fit naturally to heterogeneous environments, but they can also improve some aspects as network load or latency. Despite its overall benefits, massive use of mobile agents is restricted by security issues [8, 7]. Two main entities are considered in this scenario: the mobile agent and the host. These are the cases that can be found [1]: (1) the protection of the host against the attacks of the agents. Most of these attacks can be detected or avoided by using sandboxing techniques and a proper access control; and (2) the protection of the agent against the attacks of the host. This kind of attacks is known as the problem of the malicious hosts, and it is considered the most difficult security problem to solve in mobile agent systems [5, 8].

The aim of this paper is introducing a usable agent protection system based on dissuading the malicious hosts. This can be achieved by means of: (1) an effective attack detection mechanism. This paper explains how to use the Mobile Agent Watermarking (MAW) approach [4] to detect manipulation attacks; and (2) a punishment system. This paper also introduces how to use MAW to punish the malicious hosts by means of a Trusted Third Party (TTP), the Host Revocation Authority (HoRA) [3].

This paper is organized as follows: Section 2 resumes the state-of-the-art related solutions to solve the problem of the malicious hosts; Section 3 explains how to detect attacks using the Mobile Agent Watermarking approach; Section 4 details the HoRA functionalities, and the way in which it can be used to punish malicious hosts; finally, some conclusions can be found in Section 5.

2 Malicious Hosts

Current protection techniques can be divided in two categories: attack detection approaches, that permit the origin host to know if its agent was tampered during the execution; and attack avoidance approaches, whose aim is avoiding the attacks before they happen. Detection techniques are not useful for services in which the benefits for tampering the agent are greater than the possible punishment. In those cases, avoidance techniques are more useful. Unfortunately, there is no current approach that avoids attacks completely.

2.1 Attack Avoidance Approaches

In [10], Ordille proposed executing the agent only in trusted hosts. However, this proposal is not useful in open networks because there are few trusted hosts. Yee introduces the idea of a closed tamperproof hardware subsystem [14] where agents can be executed in a secure way, but this forces each host to buy a hardware equipment and to consider the hardware provider as trusted. Roth presents the idea of cooperative agents [11] that share secrets and decisions and have a disjunct itinerary. This fact makes collusion attacks difficult, but not impossible. Hohl presents obfuscation [6] as a mechanism to assure the execution integrity during a period of time, but this time depends on the capacity of analyzing the code. The use of encrypted programs [12] is proposed as the only way to give privacy and integrity to mobile code. The difficulty here is to find functions that can be executed in an encrypted way.

2.2 Attack Detection Approaches

In [9], the authors introduce the idea of replication of the agents and voting, but this can only be used as an attack detection approach in those scenarios in which the hosts have different interests to attack an agent. In [13], Vigna introduces the idea of the cryp-

tographic traces. The running agent takes traces of instructions that alter the agent's state due to external variables. The host stores the traces because their size depends on the amount of input data. If the origin host wants to verify execution, it asks for the traces and executes the agent again. If the new execution does not agree with the traces, the host is cheating. Verification is performed in case of suspicion, but how a host becomes suspicious is not explained. Additionally, the hosts must store the traces for an indefinite time.

3 Attack Detection

In [4], the authors introduced the Mobile Agent Watermarking approach, a lightweight attack detection approach that permits to verify the execution integrity of all the hosts without thinking in terms of suspicion. The running of the agent creates a data container where the watermark is transferred and the execution results are hidden. When the agent returns to the origin host, it applies a set of integrity rules to all the data containers. If a container does not fulfill the rules, this means that the corresponding host is malicious.

3.1 Watermark Embedding and Transference

The origin host embeds a watermark into the agent's code by using software watermarking techniques [2]. In each host, the execution of the marked code creates a logically-structured data container where the watermark will be transferred. The agent puts some information into the container, for example dummy data, input data, intermediate variable values, and finally the results. The agent diffuses (repeats values) and confuses (changes values) all this information into the container. The way this information is put into the container and the information itself constitutes the transferred watermark. The container also hides the execution results from malicious hosts.

3.2 Detecting Manipulation Attacks

When the agent returns home, the origin host tries to detect the attacks performed during execution. To do so, the origin host verifies that all the containers fulfill a set of integrity rules. These rules are inferred from the modifications performed over the original agent's code to embed the watermark. If a container does not fulfill the rules, this means that the corresponding host modified the mobile agent, so it is malicious. Notice that the way the origin host uses to verify the execution integrity is the same for all the hosts, but this does not mean that all the containers have the same watermark.

3.3 Advantages of MAW

MAW is a lightweight attack detection approach if it is compared to the most widely known proposal, the cryptographic traces approach [13]. These are some of the advantages of MAW regarding the use of traces:

- The size of the containers is determined by the programmer and can be little enough to send all them back to the origin host. On the contrary, the traces are not sent to the origin host because their size depends on the amount of input data.
- In MAW, the origin host can verify the execution integrity of all the hosts. With the traces, the verification is performed in case of suspicion.
- In MAW, the executing hosts do not need to store any kind of proof. On the contrary, the hosts must store the traces for an indefinite period of time.
- In MAW, the origin host has to apply the rules to the containers to verify the execution integrity. On the contrary, the origin host must ask for the traces and execute the agent again.

4 Punishing Attacks

In [3] the authors introduced a punishment based on host revocation. If a host acted maliciously once, it can attack other agents again. To avoid this kind of attacks this malicious host will not receive agents any more. The strength of this mechanism lies in dissuading the hosts because they can be revoked.

4.1 Host Revocation Authority (HoRA)

The lack of a TTP with punishment capabilities can be solved by adding a Host Revocation Authority to the mobile agent system [3]. The HoRA stores a database with all the information related to the past attacks. These are the main tasks that the HoRA performs:

- Status Checking: before sending an agent, the origin host consults the revocation information to delete all the revoked hosts from the agent's itinerary. As a result, the revoked hosts will not execute agents any more.
- Host Revocation: the HoRA is responsible for adding new malicious hosts to its database. To do so, the origin host must have proofs of the malicious behaviour of the host. In this paper we introduce some protocols that the origin host can use to revoke malicious hosts using MAW.

4.2 Host Revocation

The second task of the HoRA is managing the revocation information of its database. As the revoked hosts are not removed from the database, this task consists mainly in adding new hosts. As the HoRA is a TTP, the revocation of a host can only be performed in case there are proofs of its malicious behaviour. In this Section we integrate the detection mechanism of MAW [4], with the punishment capabilities of the HoRA [3].

We can divide the agent's lifetime in three phases:

- Agent Sending Part: in this phase the origin host sends the agent to perform its tasks.
- Proof Checking Part: in this phase the origin host uses MAW to find the malicious hosts.
- Host Revocation Part: if there are malicious hosts, the origin host starts a protocol to revoke them.

First of all, we focus on the phases assuming that privacy is not needed. In this case the process is easier because there is no encryption. Later, we study the modifications to the phases when privacy is needed. In both cases integrity of the agent's contents must be assured to consider the proofs valid. It is assumed that the origin host performed the status checking previously, and that the agent's itinerary has N hosts.

4.2.1 Notation

Before starting with the protocol details, some notation used in the message and agent passing must be introduced:

- We denote a mobile agent that moves from host x to host y as $Agent_{x\to y}()$.
- We denote a message from host x to host y as $Message_{x\to y}()$.
- We denote the signed copy of document D as $sign_{\alpha}[D]$, where α is the signing host.
- We denote the One-Way Hash Function value of document D as $H(D)$.
- We denote the encrypted document D with a symmetric algorithm and session key K_s as $S_{Ks}[D]$.
- We denote the decrypted document D with a symmetric algorithm and session key K_s as $S^{-1}_{Ks}[D]$.
- We denote the encrypted document D with a public key algorithm and private key K_{priv} as $E_{Kpriv}[D]$.
- We denote the decrypted document D with a public key algorithm and public key K_{pub} as $D_{Kpub}[D]$.

[1] It is assumed that the code is the same for all the hosts.
[2] The input data could be different for each host.

4.2.2 Privacy not Required

These are the three parts involved when privacy is not required.

Agent Sending Part In the agent sending part, the mobile agent travels from host to host executing its code. The steps that the entities involved must follow are included below:

1. The origin host (O) sends the following agent to the first host in the itinerary (Host-1):

   $$Agent_{O\to 1}(W_O)$$

   where

   $$W_O = sign_O[Code, Data_O, H(Rules)].$$

   The agent carries the code $Code$[1], some input data $Data_O$[2] and a hash value of the integrity rules. This hash is the proof that links the integrity rules with this particular execution.

2. When Host1 receives the agent, it verifies that the signature of W_O corresponds to the origin host. If so, it extracts the code and executes it. During the execution, the agent creates the data container. When the execution finishes, the Host-1 sends this agent to the following host:

   $$Agent_{1\to 2}(W_1)$$

   where

   $$W_1 = sign_1[W_O, Data_1, Container_1].$$

   $Data_1$ are the input data to the following hosts. $Container_1$ is the container where the watermarks are transferred and the execution results are hidden. Notice that the Host-1 does not have to store any proof because the agent carries the container.

3. This process is repeated for each host in the itinerary. So Host-i sends the following agent to the next host:

   $$Agent_{i\to i+1}(W_i)$$

   where

   $$W_i = sign_i[W_{i-1}, Data_i, Container_i].$$

4. Finally, the last host of the itinerary Host-N sends the following agent to the origin host:

   $$Agent_{N\to O}(W_N)$$

   where

   $$W_N = sign_N[W_{N-1}, Container_N].$$

Figure 1 shows the agent sending phase.

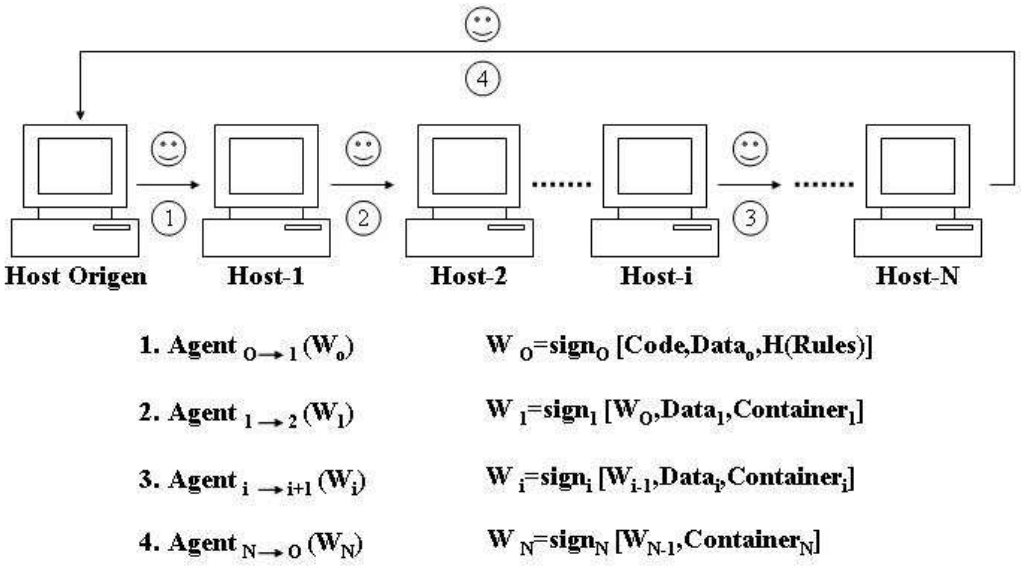


Figure 1: Agent sending part with MAW

Proof Checking Part The origin host receives the agent and verifies that all the nested signatures in W_N are valid. Once the integrity and the authentication of the data have been verified, the origin host extracts the containers of all the hosts, and it applies the integrity rules to them. If the containers fulfill the integrity rules, this means that the executing hosts are honest. On the contrary, if a container does not fulfill the rules, that is to say, if the watermark has been modified, this means that this host is malicious and hence the origin host can start a Host Revocation Protocol for MAW (HRP-M).

Host Revocation Protocol for MAW (HRP-M) Now consider that the Host-j acted maliciously modifying the agent instead of executing the code directly. So then, the execution will create a tampered container $Container_j$. These are the steps that must be followed in order to revoke Host-j:

5. The revocation process consists mainly in sending a signed message with all the proofs to the HoRA:

   $$Message_{O \rightarrow HoRA}(sign_o[W_j, Rules]).$$

   The message also contains the integrity rules.

6. The HoRA receives the revocation query and starts checking the proofs:

   - It verifies the signature of W_j to assure that the container $Container_j$ was generated by Host-j.
   - It verifies the signature of W_O to assure that the code was created by the origin host.
   - It verifies that the $Rules$ match with the hash value $H(Rules)$ that was inside W_O to verify that the code, the data, the container and the rules come from the same execution.
   - It verifies the execution integrity. The way to verify if the execution has not been tampered is by applying the integrity rules to the containers. The problem arises because these rules are not public, only the origin host knows them. So then, the HoRA needs proofs that these rules match the agent's code. This can be done by executing the agent's code (once or several times) with random input data. As the integrity rules have been inferred directly from the marked code, any container created with this code will fulfill the rules, independently from the input data. The integrity rules are valid if the containers created during these random executions fulfill the rules[3].
   - Finally, the HoRA can verify if the executing host acted maliciously by applying the integrity rules $Rules$ to the container $Container_j$. If the container fulfills the rules, this means that the Host-j is honest and hence the origin host started a revocation process with invalid proofs. In this case, the HoRA can impose a sanction or punishment to the origin host because of its dishonest attitude. On the contrary, if the container does not fulfill the rules, Host-j can be revoked. An informative message is also sent to the origin host:

     $$Message_{HoRA \rightarrow O}(sign_{HoRA}[Revoked]).$$

[3] If the random containers do not fulfill the integrity rules, this means that the origin host is acting dishonestly trying to revoke an honest host.

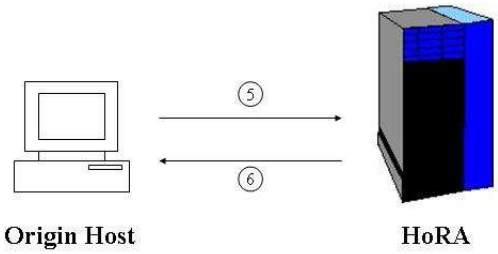


Figure 2: HRP-M

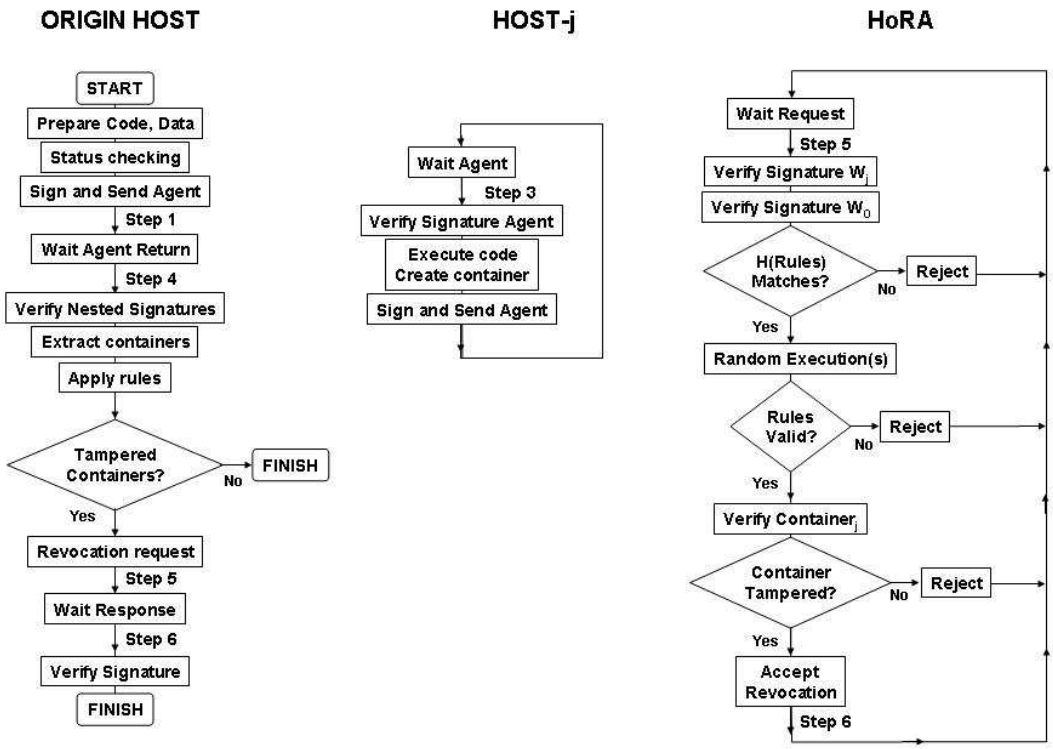


Figure 3: Flowcharts of the origin host, the Host-j and the HoRA

Figure 2 shows the message passing to revoke Host-j using HRP-M. Figure 3 shows the flowcharts of the origin host, the Host-j and the HoRA respectively.

4.2.3 Privacy Required

In some scenarios encryption must be used. However, this can adversely affect the host revocation system. The problem is making the HoRA accept proofs that have been encrypted, without revealing the private key of any of the entities. The rest of the Section introduces the message and agent passing that must be used when privacy is required, emphasizing in those aspects that differ from the previous protocol that does not require privacy.

Agent Sending Part These are the steps that must be followed:

1. Before sending the agent, the origin host generates a set of session keys to encrypt the data using symmetric encryption: K_s will be used to encrypt all the contents of the agent, and K_1, $K_2 ... K_N$ to encrypt the container of each host. So then, the following agent is sent to the first host in the itinerary (Host-1):

$$Agent_{O \to 1}(S_{Ks}[V_O], Z)$$

where The data block V_O contains basically the code, the data and the hash of the rules. All these contents are encrypted by using K_s to avoid eavesdropping attacks. The agent also carries the keys array Z that can be used by each host to obtain the session key and its own key. Z contains a hash value of V_O to certify that these keys apply to this execution in particular, in the same sense that V_O contains the hash values of all the keys K_1 to K_N. The

HoRA will need these hash values to consider the proofs as valid during revocation.

2. Host-1 receives the agent and verifies that the signature of Z corresponds to the origin host. It decrypts the first position of Z with its private key to obtain the session key K_s and its key for encrypting the container K_1:

$$[K_s, K_1] = D_{Kpriv1}[E_{Kpub1}[K_s, K_1]].$$

The Host-1 can also decrypt the code and the data with this session key:

$$V_O = S_{Ks}^{-1}[S_{Ks}[V_O]].$$

Before executing the agent, Host-1 verifies that V_O matches the hash value $H(V_O)$ inside Z, and that its key K_1 matches the hash value $H(K_1)$ inside V_O. If they do not match, this means that the keys are not the proper ones, and the host can abort the execution. If they match, the Host-1 starts the execution. The agent that will migrate to the next host is:

$$Agent_{1\rightarrow 2}(S_{Ks}[V_1], Z)$$

where

$$V_1 = sign_1[V_O, Data_1, S_{K_1}[Container_1]].$$

V_1 is encrypted with the session key K_s, so only the hosts in the itinerary can read its contents. In addition, the container is encrypted with K_1, so the origin host is the only one that can read it. The input data are not encrypted because it is assumed that they must be readable for the rest of the hosts of the itinerary.

3. This process is repeated for each host in the itinerary. So Host-i sends the following agent to the next host:

$$Agent_{i\rightarrow i+1}(S_{Ks}[V_i], Z)$$

where

$$V_i = [sign_i[V_{i-1}, Data_i, S_{Ki}[Container_i]].$$

4. Finally, the last host Host-N sends the following agent:

$$Agent_{N\rightarrow O}(S_{Ks}[V_N], Z)$$

where

$$V_N = sign_N[V_{N-1}, S_{K_N}[Container_N]].$$

Proof Checking Part When the agent returns to the origin host, it can decrypt all the contents as it knows all the keys. First of all, it decrypts V_N by using the session key K_s. Next, it verifies that all the nested signatures in V_N are valid and decrypts the containers of all the hosts. It only lasts to apply the integrity rules to all the containers to verify the execution integrity. If the containers fulfill the integrity rules, this means that the executing hosts are honest. On the contrary, if a container does not fulfill the rules, this host is malicious and hence the origin host can start a Host Revocation Protocol for MAW with Privacy (HRP-MP).

[4]Notice that Host-j acknowledged this key previously.

Host Revocation Protocol for MAW with Privacy (HRP-MP) Now consider that the Host-j acted maliciously modifying the agent. These are the steps that must be followed in order to revoke Host-j:

5. The origin host starts the revocation process sending this message:

$$Message_{O\rightarrow HoRA}(sign_o[E_{KpubHoRA}[V_j, Rules, K_j]]).$$

The message contains V_j, the integrity rules and the decryption key K_j.

6. The HoRA receives the revocation query and starts checking the proofs:

- It decrypts the contents of the message using its private key.
- It verifies that the signature of V_j corresponds to the Host-j. If so, this means that the container $Container_j$ was generated by this host.
- It verifies that the signature of V_O to assure that the agent's code was generated by the origin host.
- It verifies that the $Rules$ match with the hash value $H(Rules)$ that was inside V_O.
- It verifies that the decryption key K_j sent in the message matches the hash value $H(K_j)$ into V_O. If not, this means that the origin host sent an invalid key. In this case, the HoRA can impose a sanction or punishment to the origin host because of its dishonest attitude. On the contrary, this means that the origin host and the host used the same key[4]. The HoRA can decrypt the container $Container_j$ with this key.
- It verifies that the integrity rules are valid by executing the agent's code (once or several times) with random input data. The integrity rules are valid if the containers created during these random executions fulfill the rules.
- Finally, the HoRA verifies the execution integrity by applying the integrity rules $Rules$ to the container $Container_j$. If the container fulfills the rules, this means that the Host-j is honest. On the contrary, if the container does not fulfill the rules, Host-j can be revoked. An informative message is also sent to the origin host:

$$Message_{HoRA\rightarrow O}(sign_{HoRA}[Revoked]).$$

Figure 4 shows the flowcharts of the origin host, the Host-j and the HoRA respectively, emphasizing in those aspects that affect privacy.

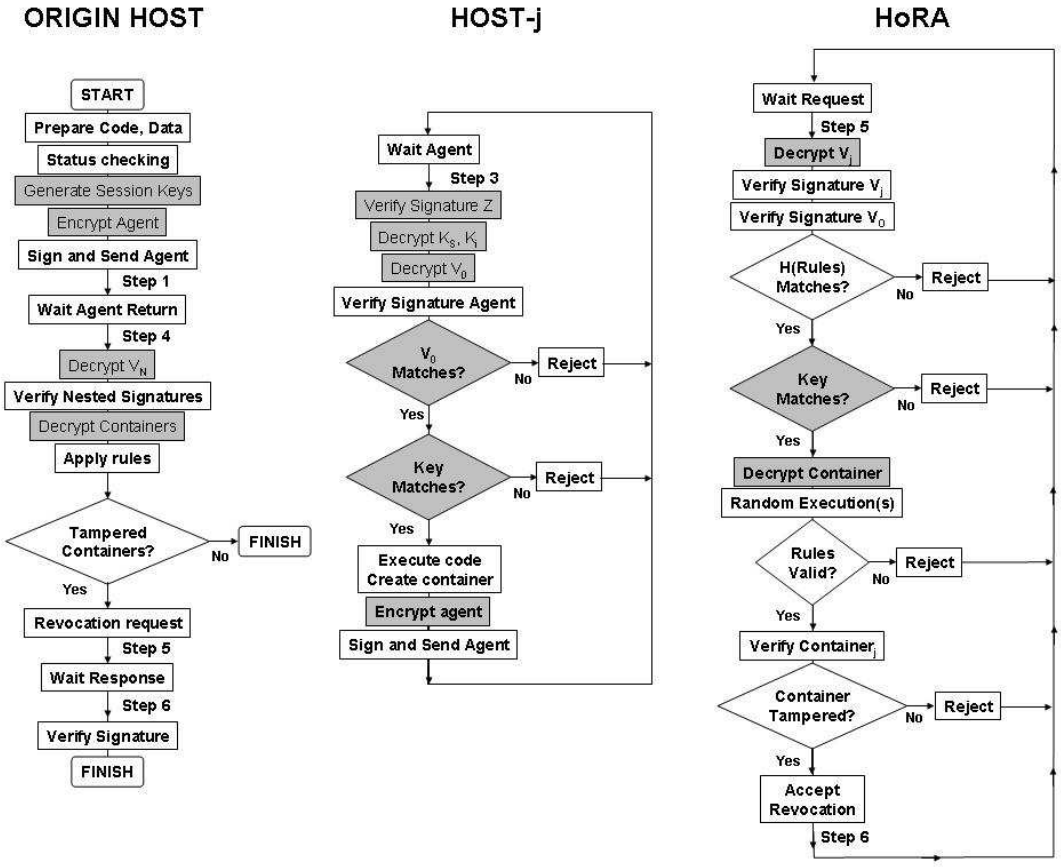


Figure 4: Flowchart of the origin host, the Host-j and the HoRA with Privacy

4.3 Attacks to the revocation protocols

These are the main attacks that the malicious hosts can perform to the revocation protocols that use MAW:

- A malicious host can try to modify any data that affect the execution: the code $Code$, the input data $Data_j$ and the container $Container_j$. Any kind of manipulation that affects the embedded watermark will be detected and hence the malicious host will be revoked.
- A malicious origin host can try to involve a honest executing host Host-i starting a revocation protocol with invalid proofs. Assuming that a honest execution creates a container that fulfills the rules, the origin host cannot involve a honest host. The origin host cannot alter the container $Container_i$ as it was signed by Host-i. The origin host can try that the container does not pass the verifications sending invalid integrity rules. In this case, the execution of the agent's code with random input data will create containers that do not fulfill these rules and the origin host will be fined for its dishonest attitude.
- In the case that privacy is required, the origin host cannot try to involve a host changing the key for the results K_i because this key must match the hash value $H(K_i)$. The Host-i cannot use a different key of K_i, because the origin host (and the HoRA) could not decrypt the results.

5 Conclusions

The results presented in this paper are twofold: presenting a usable attack detection and proving technique, and introducing a punishment mechanism to apply to the malicious host. Regarding attack detection, the Mobile Agent Watermarking approach [4] has been explained . The origin host embeds a watermark into the agent's code. In each host, this code creates a data container to transfer the watermark and to hide the results. When the agent returns home, the origin host applies a set of integrity rules that the containers must fulfill. If a container does not fulfill the rules, this means that the host has modified the agent during execution. Regarding the attack punishment, in [3] the Host Revocation Authority (HoRA) it is introduced as a TTP with punishment capabilities. The HoRA stores in

a database the information of those hosts that have been proven malicious and hence they have been revoked. Before sending an agent, each origin host consults the revocation information (status checking) in order to delete all the revoked hosts from the agent's itinerary. So then, these revoked hosts will not receive agents any more. This paper introduces some new protocols that can be used to revoke the malicious host using MAW. Different protocols have been developed depending on the privacy requirement.

Acknowledgments

This work is supported by the Spanish Research Council under the projects DISQET (CICYT TIC2002-00818) and ARPA (CICYT - TIC2003-08184-C02-02), and the European Research Council under the project UBISEC (IST-FP6 506926).

References

[1] D. Chess. Security Issues in Mobile Code Systems. In *Mobile Agents and Security*, volume 1419 of *LNCS*. Springer-Verlag, 1998.

[2] C. Collberg and C. Thomborson. Software watermarking: Models and dynamic embeddings. In *Principles of Programming Languages 1999, POPL'99*, 1999.

[3] O. Esparza, M. Soriano, J.L. Muñoz, and J. Forné. Host Revocation Authority: a Way of Protecting Mobile Agents from Malicious Hosts. In *International Conference on Web Engineering (ICWE 2003)*, volume 2722 of *LNCS*. Springer-Verlag, 2003.

[4] O. Esparza, M. Soriano, J.L. Muñoz, and J. Forné. Punishing manipulation attacks in mobile agent systems. In *IEEE Global Telecommunications Conference (Globecom 2004).*, 2004.

[5] W.M. Farmer, J.D. Guttman, and V. Swarup. Security for mobile agents: issues and requirements. In *19th National Information Systems Security Conference*, 1996.

[6] F. Hohl. Time Limited Blackbox Security: Protecting Mobile Agents From Malicious Hosts. In *Mobile Agents and Security*, volume 1419 of *LNCS*. Springer-Verlag, 1998.

[7] W. Jansen. Countermeasures for Mobile Agent Security. *Computer Communications, Special Issue on Advanced Security Techniques for Network Protection*, 2000.

[8] W. Jansen and T. Karygiannis. Mobile Agent Security. Special publication 800-19, National Institute of Standards and Technology (NIST), 1999.

[9] Y. Minsky, R. van Renesse, F. Schneider, and S.D. Stoller. Cryptographic Support for Fault-Tolerant Distributed Computing. In *Seventh ACM SIGOPS European Workshop*, 1996.

[10] J. Ordille. When agents roam, who can you trust? Technical report, Computing Science Research Center, Bell Labs, 1996.

[11] V. Roth. Mutual protection of cooperating agents. In *Secure Internet Programming: Security Issues for Mobile and Distributed Objects*, volume 1906 of *LNCS*. Springer-Verlag, 1999.

[12] T. Sander and C.F. Tschudin. Protecting mobile agents against malicious hosts. In *Mobile Agents and Security*, volume 1419 of *LNCS*. Springer-Verlag, 1998.

[13] G. Vigna. Cryptographic traces for mobile agents. In *Mobile Agents and Security*, volume 1419 of *LNCS*. Springer-Verlag, 1998.

[14] B.S. Yee. A sanctuary for mobile agents. In *DARPA workshop on foundations for secure mobile code*, 1997.

Algoritmo para la creación de una infraestructura fija virtual para gestión de claves en grandes grupos sobre MANET

Juan Hernández-Serrano, Josep Pegueroles, Miguel Soriano
Dpto. de Ingeniería Telemática - ETSI de Telecomunicaciones de Barcelona - Universidad Politécnica de Cataluña
Campus Nord - Módulo C3. 08034 - Barcelona
{jserrano,josep,soriano}@entel.upc.edu

Abstract *La Gestión de Claves de Grupo (GKM - Group Key Management) en redes inalámbricas debe abordar el problema adicional que representa la movilidad de los miembros del grupo. Las propuestas existentes de GKM sobre redes móviles se basan en esquemas centralizados y, por tanto, asumen una parte de infraestructura fija. Estas propuestas no son válidas para MANETs (Mobile Ad-hoc NETworks), que son redes sin ningún tipo de infraestructura fija y que requieren por tanto de mecanismos descentralizados. En este artículo presentamos un protocolo que permite crear una infraestructura fija virtual sobre MANET y por lo tanto utilizar los algoritmos GKM existentes. Los resultados de simulación muestran que el protocolo se adapta rápidamente a topologías dinámicas.*

1. Introducción

La tecnología MANET (*Mobile Ad-hoc NETwork*) permite la creación de sistemas autónomos de nodos móviles [4]. Las MANETs están diseñadas para operar en una gran variedad de entornos. A nivel militar, el despliegue de MANETs se puede considerar un escenario de largo alcance con cientos de nodos, heterogéneos y dinámicos. Otras MANETs pueden diseñarse para tener un alcance menor, como por ejemplo las extensiones multi-salto a la tecnología WLAN (*Wireless LAN*) 802.11. Todavía en una escala más pequeña se encuentran las aplicaciones diseñadas para redes de sensores (escasa potencia) y otros tipos de sistemas móviles.

Todos los escenarios MANET posibles mantienen una característica común: deben definir un marco de cooperación para compartir recursos y/o servicios. El marco de cooperación se extiende incluso a la seguridad, con lo que las comunicaciones seguras de grupo se encuentran con el problemas de dotar de seguridad de forma distribuida (mecanismo descentralizado) y muchas veces sin una relación previa entre los nodos.

La seguridad a nivel de grupo se consigue mediante un secreto compartido o clave de sesión. De esta forma todos los miembros del grupo pueden enviar información cifrada con la clave de sesión, autenticándose como miembros por el mero hecho de conocer la clave; y todos son capaces de descifrar la información recibida, pues conocen la clave por ser miembros del grupo.

La Gestión de Claves de Grupo (GKM - *Group Key Management*) estudia la forma de distribuir y actualizar el material criptográfico durante el tiempo de vida del grupo haciendo especial hincapié en los problemas derivados del dinamismo de los mismos [13]. La clave de sesión debe actualizarse siempre que un miembro abandone o se dé de alta en el grupo. Este proceso se denomina comúnmente con el término inglés “rekeying”. Así se consiguen los denominados servicios de confidencialidad hacia adelante y hacia atrás (Forward and Backward Secrecy - FS y BS). FS significa que ningún miembro saliente puede obtener información sobre cómo descifrar las comunicaciones de grupo posteriores a su baja. Equivalentemente, por BS se entiende que ningún miembro entrante puede descifrar comunicaciones de grupo anteriores a su alta [3].

A medida que incrementa el número de miembros del grupo la gestión de la clave de sesión se hace más difícil. La mayoría de propuestas para diseñar servicios de red escalables se basan en estructuras jerárquicas, y los algoritmos GKM que abordan esta problemática [6, 10] siguen el mismo camino. Estos protocolos dividen jerárquicamente el dominio de gestión de claves en áreas más localizadas y de más sencilla administración. El dominio queda gestionado por un distribuidor de clave de dominio o DKD (*Domain Key Distributor*), y cada área por un distribuidor de clave de área o AKD (*Area Key Distributor*). El DKD genera una clave de cifrado de datos o Clave de Dominio (DEK - Data Encryption Key) para la sesión y la distribuye a todos los AKDs. Cada AKD es responsable de distribuir de forma segura la DEK a todos sus miembros.

Teniendo en cuenta la movilidad de los nodos, los algoritmos GKM deben considerar no sólo altas y bajas de miembros, sino también roaming de miembros entre áreas sin abandonar la sesión. De esta forma el modelo de movilidad de miembros es el representado en la Fig. 1. Todos los miembros pertenecen a un mismo dominio, que se ha dividido en áreas de gestión independientes (representadas como círculos). Obsérvese que se tienen en cuenta 3 eventos: alta en el dominio o grupo seguro, baja del dominio y cambio de área. Las gestión de estos 3 eventos requiere el uso de dos tipos de algoritmos:

* **Algoritmos intra-área.** Definen la gestión de altas y bajas de los miembros del dominio dentro de cada

área. Se pueden (y deben) usar algoritmos con esquemas basados en jerarquía de claves conocidos, p.e. Logical Key Hierarchy (LKH) [7], SM-LKH [11], etc.

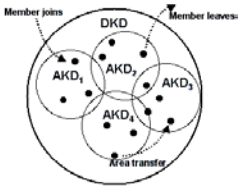


Figura 1: GKM sobre redes móviles

* **Algoritmos inter-área.** Proveen de un marco de coordinación para el intercambio de aspectos de seguridad entre áreas con las consecuentes actualizaciones de claves. Estos algoritmos definen la gestión del cambio de área de los dispositivos móviles. Ejemplos de éstos son Static Rekeying (SR), Baseline Rekeying (BR), Immediate Rekeying (IR) o Firts Entry Delayed Rekey + Periodic (FEDRP) [14].

La GKM sobre MANET debe superar un gran problema adicional además de la movilidad: la falta de una administración centralizada. Hay diversos algoritmos en el estado del arte que superan los problemas de movilidad [14], pero asumen una parte de infraestructura fija desde donde se centraliza la administración del grupo. En MANET estos algoritmos no están disponibles con lo que las claves deben ser actualizadas de forma dinámica por la propia red durante la existencia del grupo. Además, los nuevos miembros deben ser capaces de obtener y distribuir las claves del grupo de una forma ad-hoc. Para poder solventar este problema en este artículo presentamos un protocolo para crear una infraestructura fija virtual sobre un escenario MANET.

El resto del artículo se organiza como se detalla a continuación. Primero, en la sección 2, discutimos las suposiciones que se han utilizado y los objetivos de diseño, antes de detallar el protocolo propuesto en la sección 3. A continuación, en la sección 4 comparamos nuestra propuesta con otros protocolos en el estado del arte. En la sección 5, presentamos resultados de simulación, y finalmente, en la sección 6 extraemos las conclusiones y las futura líneas de trabajo de esta propuesta.

2. Suposiciones y objetivos de diseño

Describimos abajo las suposiciones que hemos tenido en cuenta en relación con los escenarios en los cuales el protocolo va a ser usado. A continuación discutimos también los objetivos de diseño de nuestra propuesta.

2.1. Suposiciones

Tal y como se define en [5], todo nodo en MANET consiste, a nivel lógico, en un router con varios dispositivos de comunicación inalámbrica. Además suponemos que todos los nodos tienen una capacidad de cómputo razonable y que pueden hacer uso de técnicas criptográficas estándar. Ejemplos de nodos serían ordenadores portátiles y PDAs modernas.

Se supone que existe un sistema de confianza por el que cada nodo da siempre información veraz de su estado. La integración con un sistema de colaboración seguro (como p.e. ePKI [9]) o uno de detección de miembros maliciosos debe ser parte de un trabajo futuro.

2.2. Objetivos de diseño

Nuestra propuesta está diseñada para su uso en grandes grupos ad-hoc seguros con nodos de movilidad limitada. El gran número de nodos en el grupo justifica el hecho de dividir de forma jerárquica el dominio de gestión de claves en áreas de menor alcance geográfico y de gestión más sencilla.

Tal y como se explica en la sección 1, las propuestas existentes de GKM sobre redes móviles asumen una parte de infraestructura fija conectada a las estaciones móviles mediante puntos de acceso, que son de hecho los AKDs de cada área. Estas propuestas están centralizadas en un DKD que también se considera parte de esa infraestructura fija. El objetivo de nuestra propuesta es proveer de un método para crear una infraestructura fija virtual en una red MANET. De esta forma se pueden extrapolar a MANETs los algoritmos de GKM móvil presentes en la literatura. El propósito del protocolo propuesto es por tanto establecer áreas a partir de un dominio de estaciones con una relación de igual a igual, y definir las relaciones entre las áreas para poder mantener un grupo seguro.

3. Propuesta de algoritmo de creación de una infraestructura fija virtual sobre MANET

Los algoritmos intra-área asumen la existencia de algunos nodos fijos que se encargan de la gestión de claves. Estos nodos fijos son de hecho puntos de acceso interconectados entre sí que actúan como AKDs. El resto de estaciones o nodos móviles se conectan a la infraestructura fija a través de estos puntos de acceso o AKDs.

En un grupo puramente MANET todos los miembros son móviles y por lo tanto no se puede asumir que tanto el DKD como los AKDs sean parte de una infraestructura fija. A continuación presentamos un protocolo para crear una infraestructura fija virtual en un escenario MANET. Como el protocolo define una infraestructura virtual que es cambiante, debe soportar cambios de AKDs y de DKD si se da el caso.

El protocolo establece áreas de gestión en un dominio de estaciones móviles con una relación de igual a igual, y define las relaciones entre áreas necesarias para mantener el grupo seguro. Las características principales de nuestro protocolo es el hecho de soportar la movilidad de todos los miembros del grupo, incluyendo aquellos que se encargan de la gestión de claves; y que

toda estación del grupo puede decidir de forma autónoma, a partir de la información de sus vecinas, si va a ser o no un gestor de área o AKD.

Vamos a definir un sistema de pesos en el que las estaciones con un peso mayor tienden a asumir el rol de AKDs. Los pesos deben ser un indicativo de parámetros como movilidad, batería, posición geográfica, etc. Como los AKDs son también estaciones móviles, se debe considerar la posibilidad de que algún AKD abandone el dominio o el grupo. En ese caso, se debe volver a ejecutar todo el algoritmo. Es por lo tanto crítico escoger un sistema de pesos de forma que el número de abandonos por parte de los AKDs sea pequeño; por ejemplo primando la baja movilidad o una buena posición geográfica. Además, es necesario remarcar que consideramos el peso de cada estación de forma única, es decir que dos estaciones no pueden tener el mismo peso.

El protocolo tiene dos fases:

- **Elección de los AKDs.** En esta fase se forman las áreas de gestión y se designa al AKD de cada área.
- **Distribución y generación de la clave de sesión o DEK.** En esta fase se define como generar la DEK para que se distribuya a los miembros del grupo.

Las altas y bajas de/en el grupo se gestionan dentro de cada área mediante protocolos GKM estándares para rekeying intra-área [7, 11]. Los cambios de área de los miembros se gestionan mediante protocolos GKM estándares de rekeying inter-área [14]. De todas formas, una re-aplicación periódica del protocolo servirá para mantener la latencia en grupos muy dinámicos. Para simplificar sugerimos el uso de SR como algoritmo de rekeying intra-área; SM-LKH como algoritmo de rekeying intra-área; y el protocolo que presentamos para generar la infraestructura fija virtual. También recomendamos un rekeying periódico para actualizar la topología de áreas de gestión y las claves en determinados intervalos de tiempo.

3.1. Elección de los AKDs

La elección de los AKDs se define de tal forma que toda estación del grupo pueda decidir de forma autónoma si va a ser un AKD o una estación ordinaria. Cada estación toma su decisión a partir de su identificación (ID) y peso, así como de los IDs y pesos de sus vecinas.

El primer paso del protocolo funciona de forma similar al propuesto en [1]. Las estaciones se asocian en clusters, estando cada cluster gestionado por una de ellas que pasa a ser un clusterhead. Los clusters definidos en el paso 1 tienen como máximo un salto de red entre cada clusterhead y cada una de las estaciones de su cluster.

Podríamos utilizar directamente los clusterheads como AKDs, pero las áreas de gestión serían demasiado pequeñas como para mejorar la eficiencia de la gestión de claves. Para superar este incoveniente, se presentan a continuación un algoritmo para crear áreas de gestión mayores mediante la unión de clusters. A continuación detallamos el algoritmo propuesto.

Algoritmo para combinar varios cluster en áreas de gestión

El algoritmo modifica la selección de clusterheads de la propuesta mencionada más arriba. Para conseguir una mayor claridad, definimos los siguientes parámetros y conceptos:

- Denominamos max_{hops} al número máximo de saltos de red entre una estación ordinaria y el AKD de su área. max_{hops} es un parámetro de diseño del sistema.
- Definimos $stations_i$ como el número de estaciones ordinarias (el clusterhead no se incluye) en el cluster i.
- C_{ij} es un cluster virtual formado por la unión de los clusters i y j. $hops_{ij}$ es el número máximo de saltos de red entre una estación ordinaria en C_{ij} y su clusterhead.
- Utilizamos el término “estación vecina” para referirnos a una estación que se encuentra a tan sólo un salto de red de la estación de referencia.
- Utilizamos el término “clusterhead vecino” para referirnos al clusterhead de un cluster unido al cluster de referencia por un salto de red. Se da en el caso de que una estación del cluster de referencia esté directamente conectada (un salto de red) a una estación de otro cluster, que denominamos “cluster vecino”.
- Para el análisis del algoritmo se hace uso del término “round”, que representa una unidad de tiempo durante la cual una estación desarrolla una acción. Por ejemplo tomar la decisión de ser un clusterhead enviando el mensaje correspondiente.

Una vez que se han elegido los clusterheads, éstos recopilan la siguiente información de sus clusterheads vecinos: peso y $stations_i$ de su área. Mediante el algoritmo propuesto, los clusterheads actuales decidirán si serán o bien el AKD de un área o bien una estación ordinaria.

Definimos dos roles: clusterhead, y estación ordinaria; y tres decisiones: petición de asociación (AR - *association request*), clusterhead, y estación ordinaria. La regla básica es que cada $clusterhead_i$ espera a publicar su decisión hasta:

- que todos sus $clusterheads_j$, $j \neq i$ vecinos con $station_j$ y peso más pequeños (en ese orden), siempre que $hops_{ij} \leq max_{hops}$, hayan decidido su rol; o
- el $clusterhead_j$, $j \neq i$ vecino con $station_j$ y peso más pequeños, siempre que $hops_{ij} \leq max_{hops}$, que no haya tomado una decisión le envía un AR.

Cada clusterhead envía un mensaje como el de (1) a sus clusterheads vecinos para publicar su decisión.

$$decision = [D||ID_{target}||T||H(D||ID_{target}||T)_{K_H^i}] \quad (1)$$

ID_{target} es la identidad del clusterhead vecino al que dirige el mensaje, aunque el mensaje se envía a todos los clusterheads vecinos. Si el mensaje no va dirigido a ningún clusterheads ID_{target} es nulo. K_H^i representa la clave que se usa para firmar el mensaje. T representa un time-stamp del round en que se envía el mensaje. El parámetro D define el tipo de mensaje:

- $D = A$, indica que el cluster con clusterhead $ID_{station}$ quiere asociarse con el cluster con clusterhead ID_{target} (AR).
- $D = CH$, indica que el clusterhead $ID_{station}$ decide que volverá a ser clusterhead. Si ID_{target} no es nulo, este mensaje es también una respuesta afirmativa a un AR de ID_{target}. Si el mensaje va dirigido a ID_{target}, una vez recibido ID_{target} y sus estaciones ordinarias asociadas pasarán a ser estaciones ordinarias asociadas a $ID_{station}$.
- $D = OS$, indica que el clusterhead $ID_{station}$ junto con sus estaciones ordinarias asociadas pasarán a ser estaciones ordinarias asociadas a ID_{target} (que en este caso no puede ser nulo). Este mensaje es siempre una respuesta afirmativa a un AR de ID_{target}.

El siguiente pseudo-código explica de forma esquemática las acciones relacionadas con el algoritmo propuesto que ejecuta cada clusterhead. Este código lo ejecuta cada clusterhead i hasta que cualquier combinación del cluster i con cada cluster vecino j produce que $hops_{ij} > max_{hops}$.

ch_i es el clusterhead actual

- Si ch_i no es el clusterhead con menor $stations_i$ y peso de sus vecinos que no ha tomado una decisión todavía

  while(true) {

  ch_j es el clusterhead con menor $stations_j$ y peso de los clusterhead vecinos de ch_i que cumple que $hops_{ij} \leq max_{hops}$ y que todavía no ha decidido su rol

  - Si ch_j es nulo
    - ch_i manda un mensaje publicando que su rol es clusterhead
      - break (sale del while);
  - Si ch_i recibe un AR ($D = A$) dirigido a si mismo ($ID_{target} = ID of\ ch_i$) de ch_j
    - Si el peso de ch_i es mayor que el de ch_j
      - ch_i envía un mensaje de decisión con $ID_{target} = ID of\ ch_j$ y $D = CH$
    - Si el peso de ch_j es mayor que el de ch_i
      - ch_i envía un mensaje de decisión con $ID_{target} = ID of\ ch_j$ y $D = OS$
    - break (sale del while);

  }

- Si ch_i es el clusterhead con menor $stations_i$ y peso de sus clusterhead vecinos que todavía no ha decidido su rol

  while(true) {

  ch_jes el clusterhead vecino con menor $stations_j$ y peso que cumple que $hops_{ij} \leq max_{hops}$.

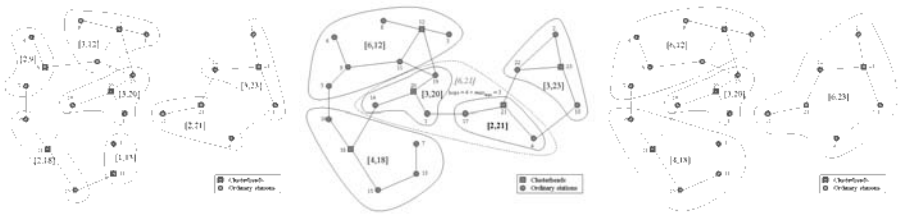
  - Si no hay ningún clusterhead vecino que cumpla que $hops_{ij} \leq max_{hops}$
    - ch_i envía un mensaje de que continua siendo clusterhead con $ID_{target} = null$ (no va a dirigido a nadie en particular), $D = CH$.
    - break (sale del while);
  - ch_i envía un mensaje de decisión AR ($D = A$) dirigido a ch_j ($ID_{target} = ID of\ ch_j$).
  - ch_i espera un mensaje de decisión de ch_j hasta que $D \neq A$ (ch_j ha decidido su rol).
  - Si ID_{target} del mensaje de decisión de ch_j es igual a ch_i(el mensaje es una respuesta afirmativa a un AR)
    - Si $D = CH$
      - ch_i manda un mensaje publicando su rol como estación ordinaria y se asocia junto con sus estaciones ordinarias asociadas al cluster de ch_j
      - break (sale del while);
    - Si $D = OS$
      - ch_i manda un mensaje publicando su rol como clusterhead
      - break (sale del while);

  }

Los clusterheads que se obtienen cuando finaliza el algoritmo anterior se corresponden con los AKDs. La generación de los clusters siguiendo criterios de número de estaciones y peso único garantiza que siempre haya algún clusterhead que comienza a publicar su decisión, y que al cabo de unos pocos rounds todos los clusterhead han tomado su decisión.

A continuación se muestra un ejemplo de cómo funciona la fase de elección de AKDs. Consideremos un grupo de 23 estaciones (Véase Fig. 2a) donde cada ID de estación se corresponde directamente a su peso. Recordemos que, por simplificación, hemos considerado que dos estaciones no puedan tener el mismo peso, con lo que aprovechamos el propio peso como identificador. Para el ejemplo decidimos $max_{hops} = 3$. El primer paso es la selección de clusterheads en [1].

Inicialmente, los nodos 18, 20 y 23 son los únicos que pueden tomar una decisión dado que no tienen ninguna estación vecina (un salto de red) con un peso mayor. Por lo tanto se declaran como clusterheads enviando un mensaje a sus vecinos. Una vez recibidos los mensajes, los nodos 15 y 16 se unen al cluster liderado por el nodo 18; los nodos 1, 14 y 19 se afilian con el nodo 20; y los nodos 2, 10 y 22 se unen al cluster liderado por el nodo 23. Cada nodo notifica su decisión de unirse a un cluster enviando el mensaje correspondiente a sus vecinos. En este punto, los nodos 12, 13 y 21 saben que sus nodos vecinos de mayor peso ya han decidido su rol (en este caso han decidido ser estaciones ordinarias) así que deciden ser clusterheads y envían el mensaje correspondiente a sus vecinos. Una vez que conocen el clusterhead de mayor peso que es vecino, los nodos 3, 8 y 11 se unen al cluster liderado por el nodo 12; el nodo 7 se une al cluster del nodo 13; y los nodos 4 y 17 se afilian con el nodo 21. El nodo 9, que estaba esperando la decisión del nodo 11, al darse cuenta de que éste no va a



(a) (b) (c)

Figura 2: Algoritmo de creación de áreas de gestión

ser clusterhead se declara como clusterhead. Finalmente, los nodos 5 y 6 pueden tomar una decisión y se afilian al nodo 9 como estaciones ordinarias. Los clusterheads elegido son los nodos cuadrados en la Fig. 2a.

Una vez que se establece el esquema de clusters, los clusters con clusterheads con menores $station_i$ y peso comienzan a asociarse para formar clusters mayores o áreas. La áreas de gestión deben estar acordes con el parámetro max_{hops} cuando se finaliza el algoritmo.

La Fig. 2a representa el esquema inicial de clusters obtenido. Para una mayor claridad identificamos los clusters como [$stations_i$,peso del clusterhead]. Inicialmente los clusters $[1,13]$, $[2,9]$ y $[2,21]$ son los únicos que puede empezar el proceso de asociación dado que no hay ningún cluster vecino con menores $station_i$ y peso. Por tanto, intentan asociarse con sus clusters vecinos de menores $station_i$ y peso con los que se cumpla que la distancia máxima en saltos de red entre las potenciales nuevas estaciones ordinarias y el clusterhead sea menor o igual que max_{hops}. Durante el primer paso, los clusters $[1,13]$ y $[2,9]$ se asocian con los clusters $[2,18]$ y $[3,12]$ respectivamente formando dos clusters nuevos $[4,18]$ y $[6,12]$ (ver Fig. 2b). El proceso de asociación de los cluster conlleva tres rounds (AR + respuesta afirmativa + asociación de las estaciones). El cluster $[2,21]$ no intenta asociarse con el cluster $[3,20]$ ya que la máxima distancia en saltos de red del cluster que formarían sería mayor que max_{hops}. De esta forma, el cluster $[2,21]$ se asocia con su siguiente cluster vecino con menores $station_i$ y peso ($[3,23]$) creando el nuevo cluster $[6,23]$. La nueva topología se representa en la Fig. 2c. En el segundo paso se para el algoritmo ya que cualquier nueva asociación de clusters rompería la condición de $max_{hops} = 3$.

Una vez terminado el nuevo esquema contiene 4 áreas de gestión cada una con su clusterhead que podemos denominar ya como AKD y con una latencia máxima garantizada de 3 saltos de red.

3.2. Generación y distribución de la clave de sesión o DEK

Una vez terminada la fase de elección de AKDs, tenemos el dominio dividido en áreas cada una liderada por su AKD, pero todavía se deben implementar las funciones del DKD, es decir se debe tratar el problema de la generación y distribución de la clave de sesión o DEK. Tenemos dos posibilidades:

1. Elegir un DKD que se encargue de estas funciones. Para la elección del DKD puede utilizarse algún algoritmo de elección de líderes, o bien elegir el AKD con mayor peso, que será en la mayoría de los casos, si la función de pesos se define adecuadamente, una de las estaciones mejor situadas y de mayor peso. Téngase en cuenta que el mecanismo de elección del líder se puede gestionar muy fácilmente puesto que está reducido a un pequeño censo electoral que son los AKDs.

2. Hacer uso de técnicas de acuerdo de claves entre los AKDs tales cómo las usadas para DPG (*Dynamic Peer Groups*) [12]. Estas técnicas derivadas del Diffie Hellman de grupo permiten la generación por colaboración de nuevas claves. De esta forma, cuando sea requerido, los AKDs generarán una nueva clave de sesión o DEK y las distribuirán a sus estaciones asociadas. Téngase en cuenta que utilizando esta técnica no se necesita elegir un DKD puesto que sus funciones quedan ahora distribuidas entre los AKDs.

El uso de la segunda técnica se adecua mejor al concepto de trato de igual a igual de nuestra propuesta, pero se comenta la primera por el hecho de una posible implementación más sencilla. Es importante remarcar también que las técnicas de acuerdo de claves no son eficientes para un acuerdo entre muchos miembros. Esto justifica que los AKDs sean los únicos miembros que forman parte del acuerdo de claves. Como el grupo de AKDs se puede considerar muy pequeño (como mucho del orden de decenas), dichas técnicas trabajan de forma eficiente con nuestra propuesta.

3.3. Alta de un nuevo miembro en el grupo o dominio

Cuando una estación quiere formar parte del grupo envía un mensaje broadcast de alta a sus estaciones vecinas. Éstas le responden con las características básicas de su área: distancia hasta el AKD, dirección y peso del mismo. Una vez obtenida la información de las diversas áreas a las que se tiene conectividad, la estación se une

Tabla 1: Comparación de protocolos de comunicación segura de grupos

Nombre del protocolo	LEAP	Protocolo definido en [8]	SR, BS, IR, FEDRP [14]	Nuestra propuesta
Grandes grupos	Sí, el dominio se divide en clusters o áreas.	No.	Sí, el dominio se divide en clusters o áreas.	Sí, el dominio se divide en clusters o áreas.
Infraestructura fija	Sí, el servidor de claves debe estar siempre disponible.	Sí, el servidor de claves debe estar siempre disponible.	Sí, tanto el DKD como los AKDs deben estar siempre disponibles.	No, aunque se crea virtualmente.
Técnicas criptográficas	Técnicas propietarias diseñadas para nodos sensores.	PKI.	No se especifica. Cualquiera soportada por los miembros (p.e. PKI).	No se especifica. Cualquiera soportada por los miembros (p.e. PKI).
Movilidad de los miembros	Se supone que todas las estaciones son estáticas.	Todas las estaciones pueden ser móviles excepto el servidor de claves.	El DKD y los AKDs deben ser estáticos, el resto de estaciones pueden ser móviles.	Todas las estaciones pueden ser móviles.
muchos-a-muchos / uno-a-muchos	Ambos	uno-a-muchos	Ambos	Ambos, aunque diseñado especialmente para muchos-a-muchos.
Fase de inicialización	El servidor de claves o estación base se establece de forma previa.	El servidor de claves se establece de forma previa.	El DKD y el AKD se establecen de forma previa.	El protocolo sirve para que cada estación decida su rol.

al grupo comunicándose con el AKD a menor distancia de red, y en caso de empate con mayor peso.

Para garantizar la confidencialidad hacia atrás (BS), se debe actualizar la clave de dominio y las claves de área que han sido comprometidas. Nuestro protocolo no define un algoritmo GKM de rekeying intra-área y se debe usar cualquiera de los existententes ([7], [11], etc). La clave de dominio se debe actualizar y re-distribuir, ya sea por el DKD en funciones o por una técnica de acuerdo de claves entre los AKDs del dominio.

3.4. Baja de un miembro del grupo o dominio

Debemos tratar dos tipos diferentes de bajas:

1. **Baja de una estación ordinaria.**

   El mecanismo es exactamente igual que para los protocolos de GKM móvil en la literatura. El AKD de su área detecta la baja e inicia la actualización de las claves de área comprometidas (si hay) y de la clave de sesión o DEK. La clave de dominio se actualiza y redistribuye, ya sea por el DKD en funciones o por una técnica de acuerdo de claves entre los AKDs del dominio.

2. **Baja de un AKD o del DKD (si existe).**

   Se debe re-ejecutar el protocolo completo para adaptarse a la nueva topología. Muchos de los AKDs actuales podrían cambiar debido a que ahora haya otras estaciones con un peso mayor, p.e. por tener menores expectativas de movilidad o mejor posición geográfica. Este hecho justifica el uso de una función de pesos que premie la baja movilidad, lo que provocará un número reducido de bajas de AKDs.

3.5. Cambio de área de un miembro del dominio

Los cambios de área se gestionan con el algoritmo de rekeying inter-área mecionados en la sección 1. Como se menciona en [14], en situaciones de frecuencia baja de altas y movilidad elevada, SR tiene la menor tasa de mensajes intra-área, de mensajes de rekeying y un reducido número de mensajes inter-área. En cambio, con mayor frecuencia de altas y movilidad más reducida, el que se comporta mejor es FEDRP. El hecho de que ambos protocolos necesiten de un rekeying periódico permite que también se reestablezca el esquema de áreas producido por nuestro protocolo. Obviamente, el periodo de rekeying debe ser sensiblemente mayor que el tiempo utilizado para establecer las áreas del dominio.

4. Comparación con otros protocolos

La Tabla 1 resume las principales características de los protocolos para comunicaciones de grupo seguras con miembros móviles y/o ad-hoc. El objetivo de la tabla es mostrar las principales diferencias entre nuestro protocolo y otros existentes en la literatura.

Nuestro protocolo se ha diseñado para grandes grupos seguros ad-hoc. El hecho de ser usado para grandes grupos justifica que divida el dominio en varias áreas de gestión; pero al contrario que en otros protocolos, los gestores de área no son fijos y pueden cambiar dinámicamente para adaptarse a la naturaleza móvil del grupo.

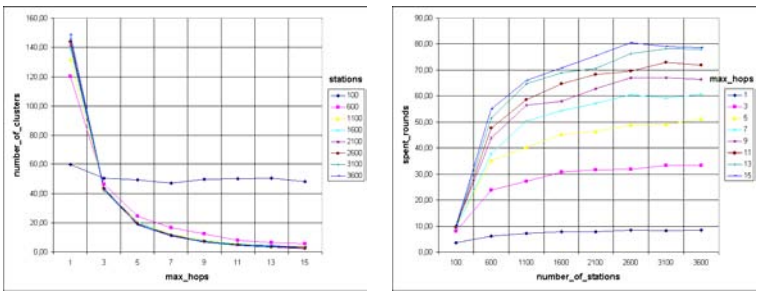
LEAP [15] también soporta la gestión de grandes grupos pero sin embargo asume como mínimo un miembro fijo que se encarga de gestionar la seguridad. Los algoritmos presentados en [14] también soportan grandes grupos pero, como LEAP, necesitan infraestructura fija que son los AKDs y el DKD.

Nótese que nuestro protocolo se ha diseñado para dotar de una infraestructura fija virtual a los algoritmos presentados en [14] y que por tanto se puede considerar como una fase de inicialización para estos algoritmos sobre un grupo puramente MANET.

5. Evaluación de eficiencia

Para poder obtener resultados simulados del protocolo presentado, hemos desarrollado un entorno de pruebas basado en las siguientes suposiciones:

- Las estaciones se sitúan de forma aleatoria dentro de una superficie rectangular de 300x300 m^2. Se toma esta disposición para simplificar el análisis, pero el simulador soporta cualquier área y disposición de las estaciones en un plano.



(a) No. final de clusters (b) Rounds empleados

Figura 3: Resultados de simulación

- La cobertura de una estación es 20 m. Esto significa que dos estaciones son vecinas si se encuentran a menos de 20 metros de distancia. Evidentemente este parámetro es modificable en el simulador.

- El protocolo ad-hoc de enrutamiento que se encuentra una capa por debajo siempre encuentra el camino más corto (en saltos de red) entre dos estaciones. El protocolo presentado no depende del protocolo de ruteo siempre y cuando pueda obtener la distancia en saltos de red entre dos estaciones. La elección del protocolo de camino más corto se basa simplemente en la facilidad de implementación dentro de nuestro simulador.

El simulador ha sido realizado con la tecnología Java2. Los resultados de simulación representados en la Fig. 3 se han obtenido a partir de la media de 10 iteraciones por cada combinación de los siguientes parámetros:

- número de estaciones en el área varía desde 100 hasta 3600 en saltos de 500 estaciones.

- el parámetro max_{hops} para cada área varía desde 1 a 15 saltos de red en saltos de 2.

El primer resultado obtenido de la simulación es que, tal y como se esperaba, el protocolo converge y no tiene probabilidad de bloqueo.

En la Fig. 3a se observa claramente como el número de áreas de gestión o clusters disminuye al incrementar el parámetro max_{hops}. Evidentemente, cuanto mayor pueda ser la distancia entre el AKD y sus estaciones asociadas, mayor número de éstas se asociarán a cada AKD, y por tanto será más pequeño el número final de áreas. Es preciso mencionar que el número de clusters no depende del número de estaciones escogido. En nuestra simulación, esto se debe al hecho de que aumentar la densidad solamente añade más estaciones que se encuentran, en su mayoría, a un sólo salto de red del clusterhead. Eso sí, los clusters o áreas tendrán un mayor número de estaciones.

La Fig. 3b representa los rounds que se han necesitado para establecer las áreas de gestión. Tal y como se preveía, son necesarios más rounds a medida que aumenta el número de clusters. Esto sucede porque el proceso de asociación de dos clusters dura al menos 3 rounds (AR + respuesta afirmativa + asociación) mientras que otros clusters están posiblemente esperando mientras toman la decisión de asociarse (AR + respuesta afirmativa).

Acotación numérica

Si consideramos una latencia máxima de 5 saltos de red ($max_{hops} = 5$) para el algoritmo GKM intra-área y 2600 estaciones (grupo grande), se puede observar que la inicialización del algoritmo (elección de AKDs) emplea una media de 46 rounds y forma 20 áreas de gestión. Tal y como se explicaba en la subsección 3.2 este número reducido de AKDs provoca que sea sencillo tanto implementar un algoritmo de elección de líder para escoger el DKD como generar la clave de dominio mediante una técnica de acuerdo de claves.

Consideremos una clave RSA como clave de firma K_H, 32 bits para el identificador único ID, 4 bits para el campo de decisión D, y 16 bits para el time-stamp del round T. Asumiendo que el mensaje de decisión de longitud máxima puede enviarse mediante un paquete de nivel físico de 512 bytes, el paquete enviado será en el peor de los casos de un tamaño máximo de 4096 bits. De las tablas en [2], se puede obtener que el caudal medio real para WLAN 802.11b (11 Mbps) con RTS/CTS, cuando hay un alto nivel de concurrencia es de 2.739 Mbps. Tomemos el caso todavía peor de que sólo fuera de 1 Mbps, entonces emplearíamos $\frac{4096}{1*1000000}s = 4{,}096ms$ para enviar un mensaje de decisión por 1 salto de red.

Ahora, en el ejemplo anterior, asumamos que se envía un paquete a una distancia de 5 saltos de red en cada uno de los 46 rounds (peor caso). El tiempo total empleado para establecer la infraestructura fija virtual sería entonces como mucho $46 * 5 * 4{,}096 = 942ms \simeq 1s$ más los tiempos de proceso. Si tenemos en cuenta que las operaciones de criptografía asimétrica emplean decenas de milisegundos en ejecutarse (RSA 1024[1]) podemos asumir un tiempo de proceso de 50 ms para cada round, siendo éste probablemente también el peor caso. De esta forma el tiempo total necesario para establecer

[1]La firma RSA1024 emplea proximadamente 15.36 ms en un PIII, 700MHz, RAM128Mb, Linux Mandrake 10

una infraestructura fija virtual con un grupo MANET de 2600 estaciones bajo las condiciones de la simulación será siempre menor que $942ms + 46 * 50ms = 3242ms$. En consecuencia, si la topología real del grupo no cambia en un tiempo superior a la cota calculada, las sucesivas topologías fijas virtuales resultantes de las ejecuciones de nuestro algoritmo permitirán el uso de GKM centralizada, ya que ésta creerá disponer en todo momento de una infraestructura fija.

6. Conclusiones y trabajo futuro

La GKM para grandes grupos de miembros móviles debe considerar los posibles cambios de área de gestión por parte de los miembros. El hecho de que deba ser escalable para grandes grupos justifica que los algoritmos en la literatura dividan el dominio de gestión en varias áreas independientes de gestión más sencilla. De estos protocolos, los más aceptados adoptan esta división asumiendo una parte de infraestructura fija (AKDs y DKDs) [14], por lo que no son útiles para grupos puramente MANET, que no tienen infraestructura.

En este artículo hemos presentado un protocolo que crea una infraestructura fija virtual que permite que los algoritmos mencionados sean válidos sobre MANETs. Nuestro protocolo asume la disponibilidad variable de cada miembro y distribuye el problema de gestión de claves entre las estaciones del grupo. Hemos presentado resultados de simulación de nuestra propuesta que muestran tiempos de inicialización, latencia máxima en saltos de red, y número de área de gestión creadas bajo diferentes condiciones de contorno. Los resultados reflejan como nuestro protocolo se adapta rápidamente a la topología cambiante de grandes grupos MANET.

Como trabajo futuro debe incluirse una evaluación de la propuesta con diferentes modelos de movilidad de los miembros y una comparación de eficiencia de nuestro protocolo con los diferentes protocolos de rekeying inter-área para ver que combinación se comporta mejor para determinados escenarios.

Agradecimientos

Este trabajo ha sido posible gracias a los proyectos DISQET (CICYT - TIC2002-00818) y UBISEC (IST-FP6 506926).

Referencias

[1] S. Basagni, K. Herrin, D. Bruschi, and E. Rosti. Secure pebblenets. In *Proceedings of the 2001 ACM Iternational Symposium on Mobile Ad Hoc Networking & Computing, MobiHoc*, pages 156–163, Long Beach, CA, October 2001.

[2] Stefano Basagni, Marco Conti, Silvia Giordano, and Ivan Stojmenovic. *Mobile ad hoc networking*, chapter 3, page 94. Wiley-Interscience. John Wiley & sons, Inc., 2004. ISBN 0-471-37313-3.

[3] Ran Canetti, Juan Garay, Gene Itkis, Daniele Micciancio, Moni Naor, and Benny Pinkas. Multicast security: A taxonomy and some efficient constructions. In *INFOCOM 99. Eighteenth annual joint-conference of the IEEE computer and communications societies*, volume 2, pages 708–716, 1999.

[4] M.S. Corson, J.P. Macker, and G.H. Cirincione. Internet-based mobile ad hoc networking. *Internet Computing, IEEE*, 3(4):63 – 70, July 1999.

[5] S. Corson and J. Macker. Mobile ad hoc networking (MANET): routing protocol performance issues and evaluation considerations. RFC 2501, 1999.

[6] T. Hardjono, B. Cain, and N. Doraswamy. A Framework for Group Key Management for Multicast Security. Internet Draft <draft-ietf-ipsec-gkmframework-03.txt>, August 2000. Work in progress.

[7] Harney and Harder. Logical Key Hierarchy Protocol (LKH). Internet Draft, 1999. Harney-sparta-lkhp-sec-00.

[8] T. Kaya, G. Lin, G. Noubir, and A. Yilmaz. Secure multicast groups an ad hoc networks. *ACM Workshop on Security of Ad Hoc and Sensor Networks*, 2003.

[9] J. Lopez, J.A. Montenegro, J.L. Vivas, E. Okamoto, and E. Dawson. Specification and Design of Security Services: The case of Advanced Authorization and Authentication. *Computers Standards and Interfaces*.

[10] Suvo Mitra. Iolus: a framework for scalable secure multicasting. In *SIGCOMM'97*, pages 277–288. ACM, September 1997. Cannes, France.

[11] Josep Pegueroles, Wang-Bin, Miquel Soriano, and Francisco Rico-Novella. Group rekeying algorithm using pseudo-random functions and modular reduction. *Grid and Cooperative Computing (GCC). Lecture Notes in Computer Science*, 3032:875–882, 2004. ISSN 0302-9743.

[12] Michael Steiner, Gene Tsudik, and Michael Waidner. Key Agreement in Dynamic Peer Groups. *IEEE Transactions on Parallel and Distributed Systems*, 2000.

[13] D. Wallner, E. Harder, and R. Agee. Key management for multicast: issues and architectures. RFC 2627, 1998.

[14] C. Zhang, B. DeCleene, J. Kurose, and D. Towsley. Comparison of inter-area rekeying algorithms for secure wireless group communications. *Performance Evaluation*, 49(1/4):1–20, September 2002.

[15] S. Zhu, S. Setia, and S. Jajodia. LEAP: efficient security mechanisms for large-scale distributed sensor networks. *ACM CSS*, 2003.

Una arquitectura AAA basada en SAML y XACML para la provisión de servicios de control de acceso a la red

Gabriel López, Antonio F. Gómez, Rafael Marín
Departamento de Ingeniería de la Información y las Comunicaciones
Universidad de Murcia
E-mail {gabilm, skarmeta, rafa}@dif.um.es

Oscar Cánovas
Departamento de Ingeniería y Tecnología de Computadores
Universidad de Murcia
email ocanovas@ditec.um.es

Abstract. *Network access control mechanisms constitute an increasingly needed service, when communications are becoming more and more ubiquitous thanks to some technologies such as wireless networks or Mobile IP. This paper presents a particular scenario where access rules are based not only on the identity of the users, but also on authorization data related to those users. In order to accomplish this general goal, it will be necessary to add to the traditional systems specific services for authentication and authorization, and also entities able to manage the information related to identity, roles and permissions. Network access will be based on the 802.1X framework and the AAA architecture, as they constitute the basis for most of the existing proposals for limiting the access to a restricted network. Those proposals will be extended using an authorization infrastructure based on SAML, the RBAC model, and XACML as the language for expressing authorization policies.*

1 Introducción

Hoy en día, se pueden utilizar diferentes propuestas de autorización en escenarios de aplicación basados en el modelo RBAC (Role Based Access Control), por ejemplo SPKI [10], X.509 AC [12] o SAML [16]. De hecho, existen varios proyectos que hacen uso de estas tecnologías, como Liberty [3] o Shibboleth [7], los cuales han extendido algunos mecanismos de autorización bien conocidos para ofrecer seguridad en entornos Web.

Este documento presenta una aproximación al control de acceso a la red basada en certificados de identidad X.509 y en atributos de autorización. Esta aproximación presenta algunos de los retos derivados de la integración de sistemas de autenticación existentes, y un sistema de autorización flexible, escalable y manejable. Nuestra propuesta está basada en los estándares SAML y XACML [13], que serán usados para expresar políticas de control de acceso basadas en atributos, sentencias y protocolos de autorización. Esta propuesta de autorización está principalmente basada en la definición de políticas de control de acceso que incluyen el conjunto de usuarios pertenecientes a diferentes dominios que podrán ser asignados a diferentes roles para obtener acceso a la red en un proveedor de servicios, bajo determinadas circunstancias. Estas políticas son los elementos centrales de nuestro sistema, y requieren la presencia de entidades responsables de su gestión. Nuestro punto de partida es un escenario de acceso a la red basado en una arquitectura AAA (Authentication, Authorization and Accounting) [9].

2 Requerimientos

Añadir un mecanismo de autorización basado en atributos a un servicio de acceso a la red implica un conjunto de requerimientos que deben ser cumplidos por cada componente de nuestro sistema.

Por un lado, la infraestructura de red debería ofrecer mecanismos de acceso estándar, los cuales también deberían ser extensibles para incorporar mecanismos de autorización. Es más, estos mecanismos de acceso deberían imponer requerimientos mínimos para los usuarios finales.

Por otro lado, la propuesta de autorización debería claramente poder adaptarse a una infraestructura de red basada en una entidad que gestiona todas las solicitudes de acceso. La propuesta debe tener en cuenta los diferentes requerimientos impuestos por los usuarios finales para intercambiar credenciales de autorización, y también debe ser lo suficientemente flexible para poder ser integrada en los protocolos existentes usados para ofrecer acceso a la red. Además, debe ofrecer soporte para escenarios inter-dominio, donde la información de autorización debería ser intercambiada entre diferentes dominios administrativos.

Finalmente, en estos escenarios inter-dominio, será necesario definir algún tipo de acuerdo entre los dominios participantes, relacionado con la gestión de los diferentes usuarios, servicios de red ofrecidos por cada organización, y otros aspectos como reconocimiento de autoridad.

2.1 Requerimientos de red

Servidor AAA

La arquitectura AAA define un elemento central que debe estar presente en cada dominio, el servidor AAA. Éste es responsable de recibir y procesar las solicitudes de autenticación, autorización y contabilidad relacionadas con los usuarios finales. En nuestra solución, la autenticación es una etapa obligatoria, aunque está fuera del alcance de este artículo como integrar el proceso de autenticación con algún sistema de gestión de identidad, como una PKI (Public Key Infrastructure). Por otro lado, la contabilidad representa una línea futura que será planteada en trabajos posteriores. De hecho, el estándar AAA no impone que un servidor AAA deba ofrecer soporte para estos tres tipos de operaciones, siendo cualquiera de ellos opcional.

En relación a la autorización, son necesarios requerimientos adicionales: el servidor AAA debe tener un módulo (módulo específico de aplicación o ASM) que deberá ser responsable de gestionar los atributos de autorización y otro para la toma de decisiones de autorización.

Finalmente, es importante mencionar que el grupo de trabajo AAA ha definido también mecanismos de transporte que serán usados para el intercambio de datos de autorización en escenarios inter-dominio.

Tecnologías de acceso a la red

En un escenario de acceso a la red, es necesario ofrecer a los usuarios algún tipo de mecanismo para consultar a un Punto de Acceso a la Red (NAP) si pueden o no acceder a esta. Este mecanismo debería ofrecer un alto grado de seguridad, y ser extensible para soportar autorización. 802.1X [15] y PANA [14] son dos propuestas diferentes que pueden usarse como tecnologías de acceso a la red. Mientras que 802.1X es un estándar de facto que puede encontrarse en la mayoría de las redes actuales, especialmente en redes inalámbricas, PANA es un prometedor trabajo en progreso. Aunque cualquiera de estas soluciones se podría utilizar en nuestro sistema, este trabajo está basado en 802.1X como la principal tecnología usada para comunicar a los usuarios finales y el NAP. 802.1X fue diseñada para ser fácilmente integrable con un protocolo capaz de intercambiar datos genéricos de autorización; este protocolo es conocido como EAP (Extensible Authentication Protocol) [11] y soporta diferentes métodos de autenticación, llamados métodos EAP. Como veremos más adelante, hay algunos métodos EAP basados en túneles (por ejemplo PEAP [3]) que pueden utilizarse para transportar información genérica, en particular, datos de autorización. Para intercambiar estos datos de autorización se requerirán extensiones a estos métodos basados en túneles.

Transporte de datos de autorización

La solución propuesta requiere un protocolo capaz de transportar solicitudes de autenticación, autorización y contabilidad desde cualquier servicio, por ejemplo un NAP, hasta el servidor AAA para ser contestadas. Del mismo modo que para las tecnologías de acceso a la red, podemos encontrar diferentes propuestas que cumplen este requerimiento, como RADIUS [17] o DIAMETER [5]. Ambas soluciones ofrecen mecanismos para intercambiar paquetes EAP entre NAPs y servidores AAA. Sin embargo, mientras que RADIUS es el estándar más ampliamente usado, DIAMETER ofrece un alto grado de flexibilidad y puede ser usado más eficientemente para cumplir los requerimientos de nuestro escenario de aplicación. Hay varias razones por las que hemos basado nuestro trabajo en DIAMETER. RADIUS, entre otras razones ampliamente detalladas en [11], tiene ciertos problemas para soportar la movilidad y el roaming.

2.2 Requerimientos de Autorización

Especificación de autorización

Hoy en día existen diferentes propuestas para representar y gestionar sentencias de autorización. Los Certificados de atributos X.509 definen una extensión al estándar X.509 que puede utilizarse para enlazar cualquier tipo de atributo a una entidad, y propuestas como SPKI y SAML ofrecen también sentencias que pueden ser empleadas para expresar, no sólo atributos, sino también pruebas de autenticación y decisiones de autorización.

El escenario de aplicación propuesto requiere una especificación de autorización que sea estándar, ampliamente aceptada, y válida para sistemas actuales. Por este motivo, SPKI no constituye una aproximación válida, ya que, a pesar de ser una solución flexible y válida; no está ampliamente desarrollada o aceptada. Por otro lado, el escenario está basado tanto en atributos de autorización, por ejemplo roles de usuario, como en decisiones de autorización y, adicionalmente, también se necesita un mecanismo para expresar consultas y respuestas de autorización. Este requerimiento no lo cumple X.509 ACs, ya que este tipo de certificados sólo puede utilizarse para expresar atributos.

Sin embargo, SAML (estándar basado en XML) ofrece una solución flexible que está siendo cada vez más empleada en entornos Web [5][7]. Además, SAML ofrece mecanismos de transporte para intercambiar datos de autorización entre las diferentes entidades que componen el sistema. Como veremos, las sentencias SAML se intercambian entre los diferentes elementos, incluso en escenarios inter-dominio, usando DIAMETER o EAP, lo que requerirá la extensión de estos.

Políticas de Autorización

Cuando un usuario final solicita un acceso a la red, nuestro sistema tiene que obtener una decisión basada en los atributos relacionados con el usuario. De este modo, es necesario expresar de algún modo el conjunto de privilegios relacionados con los usuarios que tienen asignados atributos concretos. El documento que contiene este tipo de reglas será referenciado aquí como *Resource Access Policy*.

Por otro lado, las reglas de asignación de roles, es decir, qué usuarios pueden obtener qué atributos y bajo qué condiciones, deben también ser expresadas en un documento de política, que será usado posteriormente para crear sentencias de atributos SAML. Este documento será referenciado como *Role Assignment Policy*

Para representar estas políticas pueden encontrarse en la literatura varias alternativas. En primer lugar, algunos sistemas han desarrollado su propia especificación o lenguaje de política [8]. En segundo lugar, otros sistemas se basan en XML para definir un nuevo esquema XML para expresar sus propias políticas de autorización, como Akenti [2] o PERMIS [18]. Finalmente, podemos también encontrar algunos trabajos que hacen uso de XACML, estándar basado en XML, para representar políticas de control de acceso, ya que este estándar *de facto* fue diseñado específicamente para este propósito.

El sistema propuesto hace uso de XACML, no solo para expresar estas políticas, sino también para codificar las consultas y respuestas de autorización, que pueden ser fácilmente integrables con sentencias SAML. Adicionalmente, el uso de políticas implica la utilización de un motor de autorización capaz de procesar asignación y jerarquías de roles, privilegios de autorización, recursos y obligaciones.

2.3 Requerimientos Inter-dominio

Cuando un usuario final perteneciente a un dominio particular está solicitando un acceso a la red en un dominio distinto a su dominio origen, el dominio destino debe ser capaz de obtener los atributos del usuario desde su dominio origen. Más aún, la política de acceso a recursos definida en el dominio destino debería saber como interpretar estos atributos, ya que la autorización vendrá basada en estos. Este escenario inter-dominio requiere un acuerdo a nivel de servicios o Service Level Agreement (SLA) entre los dominios participantes. Este SLA debería expresar el modo en que ambos dominios gestionarán la autenticación y autorización de sus usuarios.

3 Elementos de la Arquitectura

En el escenario de aplicación propuesto, cada usuario final pertenece a un dominio origen, donde posee un conjunto de atributos según el papel que juega. Cuando el usuario solicita acceso a la red en otro dominio, la solicitud es recogida por el servidor AAA, el cual genera una consulta para obtener los atributos asignados al usuario por la autoridad responsable de gestionarlos. Finalmente, una vez recogidos estos atributos, el servidor AAA envía una consulta a un PDP (Policy Decisión Point), el cual responde indicando si los atributos del usuario satisfacen la política de control de acceso del dominio. Además, esta política puede establecer el conjunto de obligaciones derivadas de esta decisión, por ejemplo atributos relacionados con QoS, opciones de seguridad, etc. Como se verá, este esquema general puede implantarse en un escenario mono o multi-dominio. La Figura 1 muestra los principales componentes de este escenario.

- *Usuario Final*: Es la entidad que solicita el acceso a la red. El usuario final, perteneciente a un dominio origen donde adopta uno o varios roles, tratará de obtener acceso a la red haciendo uso de éstos. En nuestro escenario, el usuario debe tener una identidad válida, por ejemplo un certificado de identidad X.509 emitido por una CA relacionada con el dominio origen, y algunos atributos que serán usados para determinar sus derechos de acceso. Inicialmente, estará situado en una red restrictiva, por ejemplo mediante una VLAN (Virtual LAN) específica.
- *Servidor AAA*: Este servidor se utiliza para gestionar el acceso a la propia red, haciendo uso del protocolo DIAMETER como mecanismo de transporte. Cada servidor AAA contiene dos módulos ASM que son responsables de negociar la generación de sentencias SAML (Source Authority), y las decisiones de autorización (PDPs).
- *Source Authority (SA)*: Este ASM gestiona la asignación de roles a usuarios. La SA recibirá peticiones, siempre a través del servidor AAA, desde dos entidades diferentes. Por un lado, usando el modelo Push, el usuario tiene que contactar con la SA para obtener sus roles, antes de solicitar el acceso a la red. Por otro lado,

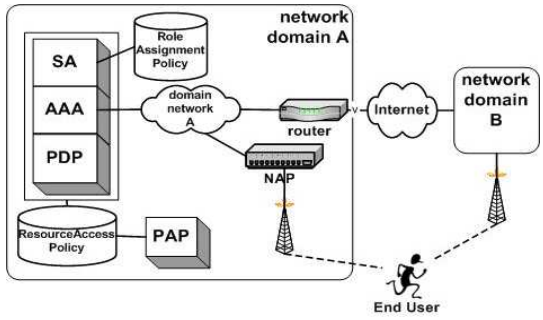


Figura 1: Elementos de la arquitectura

usando el modelo Pull, el servidor AAA local o remoto solicitará a la SA los atributos del

usuario. Cada SA tiene su propia política de asignación de roles, expresada en XACML.

- *Role Assignment Policy*: Esta política contiene las reglas que rigen la asignación de roles a usuarios. Incluye sentencias como: "en el dominio origen *Source*, el conjunto de roles *R1, R2,...Rn* puede ser asignado a los usuarios que contengan el sub-árbol X.500: *o=org,c=ES*, durante el periodo de validez *V*".
- *PDP (Policy Decisión Point)*: Este módulo ASM es el responsable de generar las sentencias relacionadas con las decisiones de autorización. Además, este elemento interactúa con el repositorio de políticas, donde está almacenada la política de acceso a los recursos. En un entorno RBAC, un PDP tiene que obtener el(los) rol(es) asignado(s) al usuario, ya que la política de control de acceso se expresa en términos de roles. Finalmente, el PDP generará una sentencia de decisión de autorización según el conjunto de evidencias presentadas.
- *Resource Access Policy*: Contiene qué dominios pueden obtener acceso a qué recursos, según los roles previamente asignados, además de las obligaciones derivadas de esta decisión. La política debería contener sentencias del tipo: "los usuarios pertenecientes al dominio origen *Source*, y jugando el rol *R1*, tendrá acceso a la red *N1* con una *QoS1*". Esta política también se expresa en XACML (los detalles sobre las políticas están fuera del ámbito de este trabajo y serán presentados en otro trabajo).
- *PAP (Policy Administration Point)*: Esta entidad define, firma y publica la política de acceso a los recursos, *Resource Access Policy*.
- *Network Access Point (NAP)*: Este elemento tiene dos funciones principales: en primer lugar, debe reenviar las solicitudes del cliente al servidor AAA apropiado del dominio destino; en segundo lugar, una vez el servidor AAA ha obtenido la decisión de autorización, el NAP obtiene y aplica las propiedades de la conexión de red definidas en la decisión. El NAP puede ser cableado o no, y debe soportar los protocolos EAPOL y DIAMETER.

4 Alternativas de Diseño

Las interacciones entre los diferentes componentes descritos en la sección previa dependen de los requerimientos impuesto por el usuario para obtener el acceso a la red. Por un lado, el usuario final puede seguir un modelo Pull, el cual requiere la mínima sobrecarga y es más apropiado para terminales limitados, como PDAs o teléfonos móviles. De este modo, todas las tareas de autorización son realizadas internamente. Por otro lado, siguiendo un modelo Push, el usuario puede presentar un conjunto de atributos, siguiendo su política de revelación. El modelo Push implica cierto soporte para la selección y transporte de los atributos, lo que supone para el usuario final una solución más intrusiva.

Sin embargo, las dos alternativas hacen uso del mismo proceso de autenticación. En primer lugar, el usuario conecta su dispositivo en un puerto disponible, o se asocia a un punto de acceso inalámbrico. En ambos casos, el NAP será configurado con 802.1X [15] para realizar el proceso de autenticación. Por lo tanto, el siguiente paso es el intercambio de paquetes EAP entre el usuario y el servidor AAA. Este servidor forzará el uso de EAP-TLS [1] para autenticar al usuario haciendo uso de su certificado de identidad. Ya que es razonable suponer que los dos dominios serán confiables, el proceso de autenticación será delegado. Una vez el usuario haya sido autenticado, la secuencia de mensajes dependerá de la alternativa seleccionada.

4.1 Alternativa de diseño 1: modelo Pull

La primera alternativa de diseño ofrece al usuario una conexión autenticada y autorizada, pero de un modo totalmente transparente, ya que la gestión de la autorización será realizada mediante un modelo pull, es decir, el PDP recuperará toda la información necesaria para tomar la decisión usando la infraestructura AAA. Una posible desventaja es que el usuario no será capaz de seleccionar el conjunto de propiedades que deben ser satisfechas por la conexión (que será determinado por la política). La mayor ventaja es que el software 802.1X usado por el usuario no tiene que ser modificado para ofrecer los servicios de autorización. Esta sección diferencia el modelo pull para entornos mono y multi-dominio.

Una vez el servidor AAA ha validado el certificado de identidad del usuario, debe chequear si el usuario ha sido autorizado o no para hacer uso de la red. El servidor AAA debe hacer uso de la *Source Authority* para obtener el conjunto de roles relacionados con el usuario antes de iniciar el proceso de decisión. En un escenario mono-dominio, la interacción es realizada por medio de una interfaz de programación (PI). El servidor AAA provee un *SAMLRequest* conteniendo un *AttributeQuery*. Esta consulta indica que la respuesta esperada debe ser codificada usando un *AttributeStatement*. Ésta también incluye información sobre el usuario que solicita el acceso, normalmente información obtenida del certificado del usuario, y, opcionalmente, el tipo de atributos que se esperan. Una vez el SA recibe la consulta, obtiene la información del usuario y establece, usando la política de asignación de roles, el conjunto de roles jugados por el usuario en el dominio origen. Estos roles pueden también estar basados en información obtenida desde el *AttributeQuery*, tal como una descripción del recurso que se está solicitando solicitado. De este modo, la SA podría seleccionar los roles más apropiados. A continuación, la SA generará un mensaje *SAMLResponse* firmado, conteniendo información sobre el estado de la consulta y un *AttributeStatement* con el(los) rol(es) del usuario.

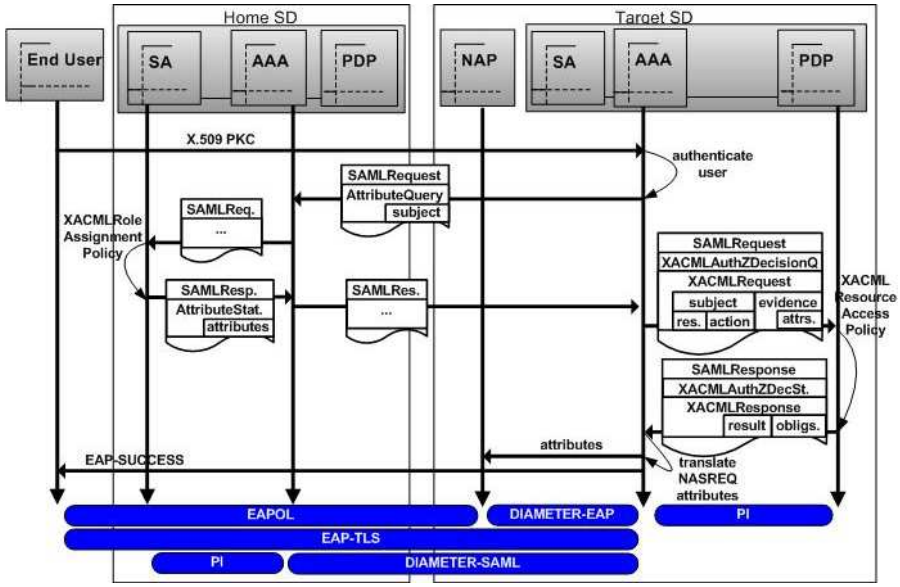


Figura 2: : Inter-domain Pull model

Sin embargo, en un escenario multi-dominio, Figura 2, cuando el usuario intenta obtener acceso a la red, el servidor AAA remoto tiene primero que descubrir cómo contactar con el servidor AAA en el dominio origen del usuario, y entonces usar algún mecanismo de transporte para intercambiar los mensajes de consulta y respuesta. La información de contacto podría estar incluida en la política emitida por el PAP, en la sección relacionada con las autoridades origen reconocidas, como en [2], o podría usarse un proceso de descubrimiento, como el definido en [7]. En relación al mecanismo de transporte, hemos definido e implementado una extensión a DIAMETER [5], DIAMETER-SAML. Esta extensión consiste en nuevos atributos DIAMETER usados para encapsular cargas SAML, siguiendo las mismas recomendaciones de otras extensiones, como DIAMETER-NAS [6] y DIAMETER-EAP [11]. Una vez el servidor AAA destino conoce al servidor AAA origen, éste envía la solicitud de consulta de atributos y espera una respuesta que contenga estos atributos.

En este punto, el servidor AAA destino tiene toda la información que necesita sobre el usuario. El próximo paso es decidir si estas evidencias satisfacen la política de acceso a recursos o no. Para realizar esto, el servidor AAA genera un *XACMLAuthorizationDecisionQuery*, la cual se envía al PDP para obtener una decisión de autorización. Esta sentencia contiene referencias al recurso siendo solicitado, el usuario que realiza la solicitud y la acción solicitada sobre el recurso. El PDP tiene que comprobar el conjunto de permisos dados a los roles relacionados con el usuario. Si el acceso solicitado es parte de estos permisos, la solicitud será concedida. Una vez la decisión ha sido obtenida, el PDP tiene que generar una sentencia *XACMLAuthorizationDecisionStatement*, que contiene el recurso destino, información sobre el usuario autorizado, los permisos que han sido concedidos, y el conjunto de obligaciones derivadas de la decisión. En nuestro escenario para el acceso a la red, el conjunto de acciones controladas se determina por algunos de los atributos definidos por la aplicación DIAMETER NAS. Finalmente, el servidor AAA obtiene la respuesta y habilita la conexión de red conforme a las obligaciones incluidas en la respuesta del PDP. Para conseguir esto, se debe realizar un intercambio de atributos DIAMETER-NAS entre el servidor AAA y el NAP.

4.1 Alternativa de diseño 2 : modelo Push

Siguiendo el modelo push, los usuarios finales son capaces de presentar sus credenciales de autorización. De nuevo, estas credenciales serán representadas usando sentencias de atributos SAML que contendrán los roles jugados por el usuario.

De acuerdo con el estándar, las sentencias SAML deben tener un tiempo de vida corto, ya que fueron diseñadas para ser usadas en entornos Single Sign On (SSO), o en escenarios de autorización Web, donde los usuarios establecen sesiones de corta duración con el recurso accedido. La ausencia de mecanismos de revocación para sentencias SAML, y su corto periodo de uso recomendado, sugieren que los documentos SAML no deben ser almacenados por entidades intermedias, como repositorios de certificados, etc. Nosotros afrontamos este inconveniente ofreciendo diferentes alternativas para obtener sentencias de atributos SAML temporalmente válidas que pueden ser consideras como confiables. Como veremos, los usuarios finales son capaces de obtener sus atributos no solo desde su dominio origen, sino también, cuando se conectan a una red destino restringida.

Una vez el usuario ha recibido toda la información que necesita, inicia un proceso de autenticación 802.1X con el servidor AAA destino. En este caso, se hará uso de PEAP (Protected EAP), que define como establecer un canal TLS usado para autenticar a las partes y para protegerlos mensajes (SAML y XACML) relacionados con la autorización.

Durante la etapa de autorización, ilustrada en la Figura 3, el usuario tiene que seleccionar los atributos que va a presentar como evidencias y, opcionalmente, qué tipo de servicio de red quiere obtener. Así, el software cliente genera un mensaje *SAMLRequest* incluyendo una *SAMLAuthorizationDecisionQuery* con los siguientes elementos: el usuario que solicita el acceso a la red, el identificador del recurso, la acción y los atributos que serán usados como evidencias. Este mensaje será enviado usando PEAP al servidor AAA, y reenviado al PDP. Una vez el PDP obtiene la solicitud, el resto del proceso sigue el mismo procedimiento que se explicó para el modelo pull.

Este diseño impone dos requerimientos: primero, el usuario debe ser capaz de recuperar sus atributos usando el mismo mecanismo independientemente de que se encuentre en el dominio origen o destino. Segundo, se debería hacer uso de un servidor Web que ofrecería una interfaz de acceso intuitiva al servidor AAA, ya que los usuarios finales no deben tener soporte para el protocolo DIAMETER.

Como muestra la Figura 4, cuando el usuario se conecta a su dominio origen, después de una fase de autenticación, el servidor web crea una sentencia *SAMLAttributeQuery* que contiene la identidad del usuario cuyos atributos serán obtenidos. Esta consulta es enviada usando DIAMETER-SAML al servidor AAA del dominio, y según las reglas establecidas en la política de asignación de roles, se devuelve una lista de sentencias *SAMLAttributeStatement* al servidor web. De este modo, el usuario obtiene la lista de roles que puede jugar. Es importante recalcar que estas sentencias de atributos son credenciales de corta duración y, por lo tanto, deben ser usadas a continuación, por ejemplo, en un escenario similar al mostrado en la Figura 3.

La principal ventaja de esta alternativa es que ofrece al usuario final completa visibilidad y control sobre el proceso de autorización, ya que puede seleccionar el tipo de conexión, características de seguridad, calidad de servicio, etc. Por otro lado, el software usado por el cliente debe ser modificado para soportar sentencias SAML, como se puede encontrar en otras propuestas como [19].

5 Trabajo Relacionado

Los servicios Web constituyen el escenario clave donde servicios de autorización y autenticación están teniendo cada vez más aceptación. Una de las principales razones es el auge de SAML como lenguaje de especificación para este tipo de credenciales, y su uso en soluciones como los proyectos Shibboleth y Liberty Alliance.

Shibboleth define un escenario de control de acceso para entornos web, y ofrece servicios de autenticación de usuarios y autorización basada en atributos y SAML. Shibboleth no ofrece la solución más apropiada en un entorno de control de acceso a la red como el propuesto. En primer lugar, la interacción entre los componentes se realiza a través del Web, usando redirecciones HTTP entre los diferentes componentes. Sin embargo, en el escenario de aplicación propuesto, como en la mayoría de los escenarios de control de acceso existentes para redes inalámbricas, el usuario solicita acceso a la red usando un modelo estándar, como 802.1X. En segundo lugar, Shibboleth sitúa la autenticación del usuario en un servicio de autoridad de autenticación,

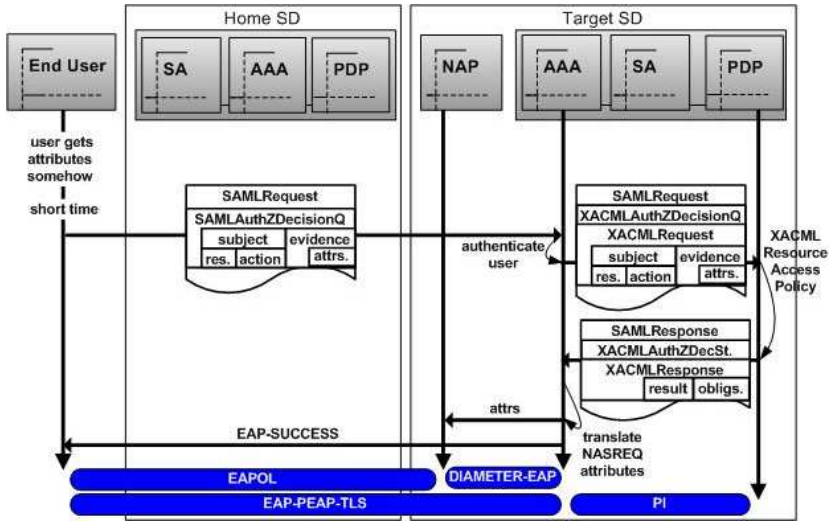


Figura 3. Inter-domain Push model

localizado en el dominio del proveedor de identidad, y basado en SAML. Por el contrario, en la solución propuesta, la autenticación se realiza de modo delegado en cada dominio, basada en una infraestructura X.509, y gestionada por un servidor AAA.

El proyecto Liberty Alliance ofrece un entorno de gestión de identidad basado en SAML a través de servicios web, aunque a diferencia de Shibboleth, el principal objetivo es ofrecer un entorno para establecer federaciones de identidad de usuarios. Liberty está basado principalmente en el intercambio de sentencias de autenticación entre autoridades y servicios, pero puede ser extendido para manejar atributos. Como Shibboleth, esta solución no es adaptable al escenario propuesto en este artículo. En primer lugar, Liberty, como Shibboleth, no soporta otro tipo de interfaz de usuario que no sea un navegador web. En segundo lugar, la especificación de Liberty ha adaptado los perfiles definidos por SAML (Liberty Artifact Profile y Liberty Browser Post Profile), haciendo más difícil la integración con otros entornos basados en el estándar. Más aún, debido a la gestión de la federación, Liberty define un amplio número de perfiles y protocolos incrementando la complejidad del sistema.

6 Conclusiones

En relación a las tecnologías involucradas, se pueden obtener varias conclusiones. En primer lugar, aunque el escenario propuesto está basado en 802.1X, este puede ser fácilmente adaptado para ser usado con otras tecnologías, por ejemplo, PANA. En segundo lugar, protocolos como EAP y DIAMETER permiten ser extendidos para transportar sentencias SAML. Finalmente, el uso de soluciones basadas en XML, SAML y XACML, para propósitos de control de acceso a la red abre un interesante campo de investigación.

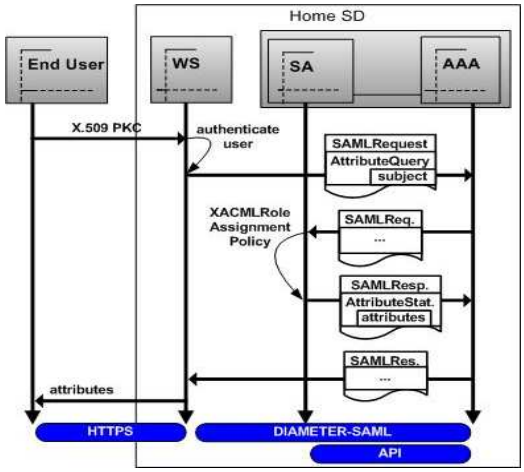


Figura 4: Recuperación de atributos desde el dominio origen

En relación a las alternativas de diseño, este documento presenta diferentes soluciones RBAC que pueden seleccionarse individualmente para desplegar el servicio de control de acceso que mejor se adapte a un entorno particular. La autorización puede realizarse de forma transparente, desde el punto de vista del usuario, usando el modelo Pull. El modelo Push sobrecarga ligeramente el sistema, pero ofrece más opciones para el usuario.

Agradecimientos

Este trabajo ha sido parcialmente financiado por los proyectos Euro6IX (IST-2001-32161) y SEINIT (IST-2002-001929).

Referencias

[1] B. Aboba, D. Simon. PPP EAP-TLS Authentication Protocol. Internet Engineering Task Force, Octubre 1999. Request for Comments (RFC) 2716.

[2] Akenti Distributed Access Control. http://www-itg.lbl.gov/Akenti/. Marzo 2005.

[3] H. Anderson, S. Josefson, G. Zorn, D. Simon, A. Palekar. Protected EAP Protocol (PEAP), 2004. IETF Draft.

[4] J. Beatty et al., Liberty Protocols and Schema Specification Version 1.1. Liberty Alliance Project. http://www.projectliberty.org/specs/archive/v1_1/liberty-architecture-protocolsschema-v1.1.pdf., Enero 2003.

[5] P. Calhoun, J. Loughney, E. Guttman, G. Zorn, J. Arkko. Diameter Base Protocol. Internet Engineering Task Force, Septiembre 2003. Request for Comments (RFC) 3588.

[6] P.R. Calhoun, G. Zorn, D. Spence, D. Mitton. Diameter Network Access Server Application, Julio 2004. IETF Draft.

[7] S. Cantor. Shibboleth Architecture. Protocols and Profiles. http://shibboleth.internet2.edu/docs/draft-mace-shibboleth-arch-protocols-01.pdf. Mayo 2004.

[8] Community Authorization Service. http://www.globus.org/security/CAS/GT3/. Marzo 2005.

[9] C. de Laat, G. Gross, L. Gommans, J. Vollbrecht, D. Spence. Generic AAA Architecture. Internet Research Task Force, Agosto 2000. Request for Comments (RFC) 2903.

[10] C. Ellison, B. Frantz, B. Lampson, R. Rivest, B. Thomas, T. Ylonen. SPKI certificate theory. Septiembre 1999. Request for Comments (RFC) 2693.

[11] P. Eronen, T. Hiller, G. Zorn. Diameter Extensible Authentication Protocol (EAP) Application, Agosto 2004. IETF Draft.

[12] S. Farrel, R. Housley. An Internet Attribute Certificate Profile for Authorization, Abril 2002, Request for Comments (RFC) 3281.

[13] S. Godik, T. Moses. OASIS eXtensible Access Control Markup Language (XACML) Version 1.0, Febrero 2003. OASIS Standard.

[14] P. Jayarama, R. López, Y. Ohba, M. Parthasarathy, A. Yegin. PANA Framework, Septiembre 2004. IETF Draft.

[15] LAN MAN Standards Committee of the IEEE Computer Society. IEEE Draft P802.1X/D11: Standard for Port based Network Access Control, Marzo 2001.

[16] E. Maler, P. Mishra, R. Philpott. Assertions and Protocol for the OASIS Security Assertion Markup Language (SAML) V1.1, Septiembre 2003. OASIS Standard.

[17] C. Rigney, A. Rubens, W. Simpson, S. Willens. Remote Authentication Dial In User Service *(RADIUS).* Internet Research Task Force, Abril 1997. Request for Comments (RFC) 2138.

[18] PERMIS. PrivilEge and Role Management Infrastructure Standards Validation. http://sec.cs.kent.ac.uk/. Marzo 2005.

[19] P. Nikander. Authorization and charging in public WLANs using FreeBSD and 802.1x. In Proceedings of the Freenix

Gestión de identidades basada en Liberty para servicios de Internet móvil

Juan Carlos Yelmo García, Jorge Ysart Álvarez de Toledo, Rubén Trapero Burgos
Departamento de Ingeniería de Sistemas Telemáticos. Universidad Politécnica de Madrid
ETSI de Telecomunicación. Ciudad Universitaria S/N
28040 – Madrid
Teléfono: 91 336 6830 Fax: 91 336 7333
E-mail: jcyelmo@dit.upm.es

Abstract. *New convergent services raise new technical and business challenges. In the case of mobile Internet services, increasing competition, a lower growth in cellular penetration rate and the need to increase subscriber ARPU are pushing for new business models and technologies supporting new session models coping with issues such as identity management, service providers federation, circles of trust, roaming among service providers, etc. This paper summarises some relevant results of a research project which has explored the use of Liberty technology for federated identity management to support new business and session models, providing a better user experience in a secure and trusted environment. For this purpose, an open source implementation of Liberty phase 1 specification has been extended to cover phase 2 and partly phase 3 and then used to support identity and session management in a prototype implementation of a location-based instant messaging service.*

1 Introducción

El mercado de la telefonía móvil es un mercado completamente maduro en España y en Europa (más del 87 % de los españoles tiene un móvil) por lo que el crecimiento de la tasa de penetración de la telefonía móvil está bajando, consolidándose a su vez el volumen de mercado de las operadoras. En este contexto, las operadoras están abocadas a buscar nuevas fuentes de ingresos, por lo que más allá de consolidar resultados económicos y reducir la tasa de rotación de clientes, deberán aprovechar las ventajas que ofrecen las nuevas tecnologías para el desarrollo y despliegue de nuevos servicios móviles tanto para el segmento residencial como el empresarial.

Por otra parte, está en marcha un proceso de convergencia tecnológica en torno al protocolo IP como tecnología de interconexión y otro de apertura y desregulación que involucra a varios actores y dominios administrativos en la provisión de servicios avanzados (redes celulares, Internet, red telefónica conmutada, terceros proveedores, etc.).

El término *Redes de Siguiente Generación* suele referirse a un enfoque de convergencia de las redes mencionadas en torno al protocolo IP como tecnología de transporte para voz, datos y multimedia. Se trata de una tecnología barata y fácil de gestionar en comparación con las redes celulares y telefónicas, aunque también presenta problemas de escalado, seguridad y garantía de calidad de servicio. Es probable que en el largo plazo las redes se construyan en torno a un núcleo de transporte basado en IPv6. En este contexto las redes actuales seguirán existiendo aunque se considerarán redes de acceso interconectadas al núcleo de red con IP como tecnología de integración (Fig. 1).

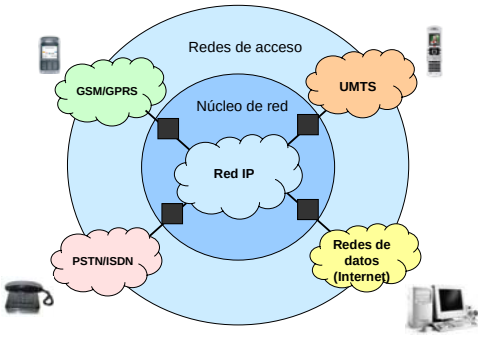


Fig. 1. Redes de siguiente generación.

La opción de convergencia total de la figura anterior es un escenario futurible, sin embargo podemos afirmar que el comienzo del proceso de convergencia ha estado marcado por el acceso a servicios de Internet móvil desde terminales celulares gracias a tecnologías como GPRS o el servicio japonés *i-mode* [1].

Las redes de siguiente generación plantean múltiples retos técnicos y de negocio. Uno de los aspectos clave en este sentido es el de los nuevos modelos de negocio que aparecen motivados por la desregulación y la creciente competencia. Estos nuevos modelos de negocio pueden implicar la intervención de varios proveedores en diferentes dominios administrativos para la provisión de un servicio. Las redes convergentes deben proporcionar mecanismos para la negociación de recursos de red y servicios entre diferentes participes en distintos dominios

administrativos. Un caso de especial interés en el contexto de este artículo es el de la **federación** entre proveedores de servicio: itinerancia de usuarios entre proveedores, federación de servicios y proveedores, federación entre proveedores heterogéneos, etc.

En este artículo se resumen algunos resultados interesantes de un trabajo de investigación sobre la gestión de identidades y los nuevos modelos de sesión y negocio en servicios avanzados de Internet móvil.

La siguiente sección describe el enfoque propuesto por la *Alianza Liberty* para gestión de identidades federadas. A continuación se describe el rol que puede jugar *Liberty* en las nuevas arquitecturas para servicios de Internet móvil. El articulo continua con un resumen de los resultados iniciales de validación del enfoque propuesto a través del proyecto FIRMA, un proyecto exploratorio desarrollado en el Departamento de Ingeniería de Sistemas Telemáticos de la Universidad Politécnica de Madrid en colaboración con la Cátedra Ericsson de la E.T.S.I. de Telecomunicación. El artículo termina con las conclusiones más relevantes del trabajo de investigación y con una breve descripción de los trabajos presentes y futuros.

2 Gestión de Identidades y tecnología Liberty

Cuando un usuario interactúa con un servicio a través de Internet, con frecuencia tiene que autenticar su identidad, porque el servicio sea de pago o tenga restringido su uso por cualquier otro criterio. Para ello, el usuario abre una cuenta con el proveedor del servicio y establece un identificador de usuario y una contraseña, junto con un conjunto de atributos que personalizan el servicio y establecen las preferencias del usuario. Se llama *Identidad de Red* al conjunto global de atributos que establecen la identidad y el perfil de un usuario en todas las cuentas abiertas por el usuario en la red. Hoy día las cuentas de usuario están dispersas en múltiples sitios en Internet, donde se prestan servicios de forma aislada y sin cohesión entre las múltiples identidades y preferencias del usuario.

Esta proliferación incontrolada de cuentas sin relación entre si conlleva una mala experiencia de usuario y merma su confianza en los servicios telemáticos. Además, presenta serios problemas de escalado, es un freno a la expansión de los servicios de comercio electrónico y atención al ciudadano y plantea un considerable riesgo de fraude por robo de identidad.

Se entiende por *Gestión de Identidades* a la disciplina que trata el problema de la gestión de acceso de los usuarios a recursos de red en sus aspectos técnicos, legales y de negocio. A nivel técnico, la gestión de identidades tiene que ver con áreas como la seguridad en redes, la provisión de servicios, la gestión de clientes, el registro único de usuario y la prestación de Servicios Web.

Existen dos enfoques básicos para la gestión de identidades en servicios de red. El primero es el enfoque centralizado, donde una única entidad gestiona atributos y elementos de identificación de todos los usuarios de servicios de red y ofrece servicios de autenticación en nombre de los proveedores de servicio. Como ejemplo de este enfoque podemos citar la iniciativa *.NET Passport* de la empresa Microsoft [1]. El enfoque alternativo es el descentralizado o federado, en el que los proveedores de servicios finales o de autenticación federan sus sistemas de gestión de identidades para permitir que los usuarios naveguen entre servicios sin autenticarse de nuevo, sin poner en riesgo la privacidad de los datos de usuario o la seguridad en el acceso a los servicios.

En Septiembre de 2001 se creó la **Alianza Liberty** [2] con el propósito de elaborar un conjunto de estándares para *Gestión de Identidades Federadas*. Se trata de un consorcio de empresas, proveedores e instituciones interesadas en proporcionar estándares y directrices para gestión de identidades federadas con garantía de privacidad y seguridad para la información de *Identidad de Red* de los usuarios. La idea básica del proyecto es proporcionar un mecanismo abierto y estándar de *Registro Único* (*Single Sign-on*) que incluye autenticación descentralizada y autorización desde múltiples proveedores. El mecanismo de *Registro Único* permite a un usuario registrarse en un proveedor de servicios de gestión de identidades y que el registro se transfiera de forma transparente cuando navega hacia otros proveedores de servicio sin necesidad de autenticarse de nuevo. La infraestructura de gestión de identidades propuesta debe soportar todos los dispositivos de acceso desde los más convencionales a los más novedosos.

En el enfoque del proyecto *Liberty*, los proveedores de servicio se asocian en *Círculos de Confianza* (*Circles of Trust*) que se apoyan en la tecnología *Liberty* y en acuerdos operativos donde se definen relaciones de confianza entre los proveedores (Fig. 2). Por otra parte, los usuarios pueden federar las cuentas aisladas que tienen establecidas con diferentes proveedores dentro del *Círculo de Confianza*. En otras palabras, un *Círculo de Confianza* es la federación de un conjunto de Proveedores de Servicio (SP) y Proveedores de Identidad (IDP) que tienen relaciones de negocio basadas en la arquitectura *Liberty* y una serie de políticas y directrices que permiten que los usuarios realicen transacciones con los proveedores de forma

[1] http://www.passport.net

[2] http://www.projectliberty.org

segura y transparente. Los *Proveedores de Identidad* serían proveedores de servicio dispuestos a gestionar la identidad federada de los usuarios de los servicios del *Círculo de Confianza* y ofrecer a los proveedores de servicio incentivos comerciales para su afiliación. Un ejemplo típico sería el de una compañía aérea como proveedor de identidad de las empresas afiliadas a su programa de fidelización.

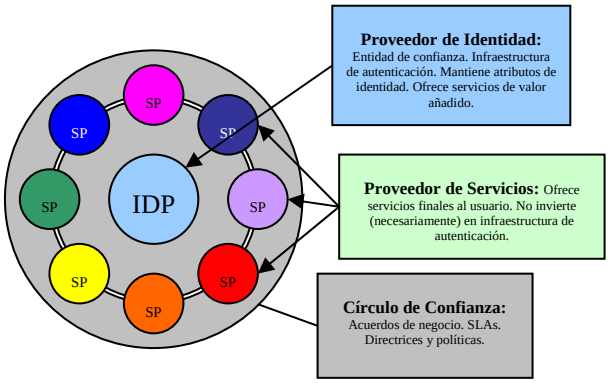


Fig. 2. Concepto de Círculo de Confianza.

El proyecto *Liberty* plantea el diseño de una arquitectura y un conjunto de protocolos que deben proporcionar soporte a la gestión de identidades federadas. El desarrollo del conjunto de estándares y recomendaciones se ha estructurado en tres fases (ver Fig. 3):

- ✓ **Fase 1**. *Identity Federation Framework* (ID-FF): Conjunto de protocolos que permiten llevar a cabo la Federación de Identidades, el Registro Único de Entrada (*Single Sign-On*), la utilización de Pseudónimos globales y el Registro Único de Salida (*Single Logout*).

- ✓ **Fase 2**. *Identity Web Services Framework* (ID-WSF): Marco de trabajo para la construcción de Servicios Web basados en Identidad: descripción y descubrimiento de servicios, autenticación, acceso a atributos compartidos, interacción con el usuario para solicitud de directrices de privacidad, etc.

- ✓ **Fase 3**. *Identity Services Interface Specifications* (ID-SIS). Definición de unos servicios específicos basados en identidad que hacen uso del marco de trabajo de la Fase 2: perfil personal, perfil de empleado, servicios de presencia, localización, alerta, calendario, monedero, etc.

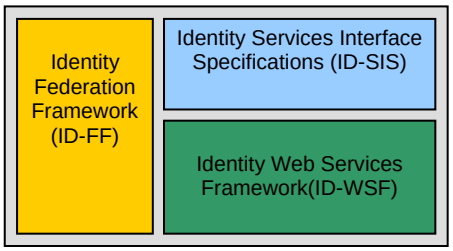


Fig. 3. Estructura de las especificaciones Liberty.

3 Liberty en el contexto de los servicios de Internet móvil

Entendemos aquí los servicios de Internet móvil como una adaptación de los servicios y aplicaciones de Internet al entorno de los terminales móviles con capacidad de proceso local y de conexión en red en base al protocolo IP desde redes de acceso celular de banda ancha. Se trata de servicios en parte limitados por la menor capacidad de proceso, almacenamiento y presentación de los terminales celulares, pero también más sofisticados por las capacidades de movilidad, localización y control de presencia junto con las posibilidades de aplicación de nuevos modelos de negocio innovadores o más propios de las redes privadas gestionadas por un operador que de las redes públicas como Internet. Estas características propician la aparición de nuevos servicios que pueden llegar a ser viables desde el punto de vista del negocio: servicios basados en localización, servicios P2P, servicios de pago por móvil, etc.

Una de las experiencias más citadas en este sentido es la del servicio *i-mode* [1] desplegado por NTT-DoCoMo en Japón en 1999. Se trata de un servicio técnicamente sencillo: acceso desde un micro-navegador específico a contenidos representados en una versión simplificada de HTML (cHTML). La característica más notable de *i-mode* es, sin embargo, el modelo de negocio aplicado. Se trata del acceso a contenidos proporcionados por terceros proveedores afiliados al operador que pueden cargar al usuario una (pequeña) cuota de abono. Esta cuota se añade a la factura del operador, quién por actuar de intermediario en el cobro percibe una comisión, además de facturar al usuario por el tráfico generado. El operador valida y aprueba los contenidos de los proveedores afiliados (sitios oficiales) y favorece el acceso a estos servicios pre-configurando sus direcciones en menús específicos de los terminales *i-mode* comerciales. Sin embargo, el usuario también puede acceder a servidores no oficiales sin más que conocer el correspondiente URL, aunque en este caso los contenidos no están validados y el proveedor no tiene integrado su sistema de facturación con el operador. Es el llamado modelo *semi-walled garden*.

Se trata en resumen de un modelo de negocio en que los proveedores de servicios y contenidos ven ampliada su base de clientes gracias al *marketing* gratuito proporcionado por el operador y sin necesidad de un sistema de facturación propio, los operadores incrementan su facturación por el volumen de tráfico generado y por su intermediación en el pago de las cuotas de abono y los usuarios tienen acceso directo a contenidos con un modelo de pago sencillo a través de una única factura emitida por una entidad de confianza.

El modelo de negocio de *i-mode* ha servido de inspiración para el trabajo que se presenta en este artículo. Suponemos por tanto que un operador

celular ofrece acceso a servicios y contenidos de Internet móvil proporcionados por un colectivo de proveedores de servicios con quienes tiene acuerdos de negocio para operar los servicios en el marco de un conjunto de políticas comunes y unos acuerdos de nivel de servicio (SLA). Esta descripción se corresponde en gran medida con la definición *Liberty* de Círculo de Confianza (ver sección 2).

En este artículo se presentan los resultados iniciales de una línea de investigación que explora la aplicación del modelo de gestión de identidades propuesto por la *Alianza Liberty* para soportar el modelo de negocio descrito más arriba y enriquecerlo con una tecnología que mejore la experiencia de usuario en un entorno de confianza y seguro y que sirva para soportar un modelo de gestión de sesiones en servicios móviles de nueva generación [2].

La siguiente figura muestra de forma esquemática la arquitectura general para la provisión de servicios de Internet móvil y el rol a jugar por la tecnología *Liberty* en ese contexto.

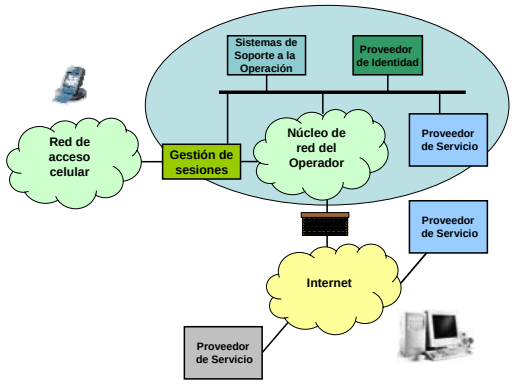


Fig. 4. Arquitectura de servicios móviles con Liberty.

4 El proyecto FIRMA

El proyecto FIRMA (*Federación de Identidades en Red para servicios Móviles basada en el proyecto Liberty Alliance*) es un proyecto de investigación que se ha centrado principalmente en el ámbito de la gestión de identidades en servicios de Internet Móvil utilizando el marco de trabajo propuesto por el proyecto *Liberty Alliance* (tanto para la gestión de identidades federadas como para el acceso a *Web Services* basados en identidad).

En los siguientes apartados se detallarán los objetivos que ha perseguido este proyecto, su estructura, las entidades y los protocolos *Liberty* que se han implementado, las herramientas utilizadas, los pasos necesarios para su integración con un servicio ya existente y un resumen de los logros y resultados obtenidos.

4.1 Objetivos del proyecto

El objetivo principal del proyecto FIRMA ha sido el desarrollo de un prototipo que permita la validación de la arquitectura y los protocolos especificados por *Liberty Alliance* para la gestión de identidades en red y su aplicación en Servicios Móviles 2.5/3G.

A grandes rasgos las facilidades que este prototipo debía cubrir eran las de federación de identidades, autenticación, registro único de entrada y salida, gestión de perfiles de usuario, acceso compartido a atributos del usuario e instanciación transparente de sesiones de acceso y servicio.

La implementación de este prototipo pasaba por diseñar y validar una arquitectura y un conjunto de componentes y servicios *middleware* reutilizables para la gestión de acceso a servicios móviles avanzados.

También se perseguía como objetivo secundario la validación del entorno de código fuente abierto publicado por *SourceID, ID-FF 1.1 Java Toolkit* [3] para el desarrollo de sistemas de gestión de identidades basados en arquitectura Liberty.

La validación del prototipo se pretendía llevar a cabo mediante la integración con un prototipo de servicio móvil avanzado ya desarrollado. En este punto hay que señalar que dicha integración se ha realizado con un *Prototipo de Mensajería Instantánea P2P basado en Localización*.

4.2 Fases del proyecto

Al igual que el proyecto *Liberty Alliance* (ver Fig. 3), este proyecto se ha estructurado en tres fases bien diferenciadas:

- **Fase 1, ID-FF**: Para la implementación de las funcionalidades *Liberty* Fase 1 el proyecto se ha apoyado en *SourceID, ID-FF 1.1 Java Toolkit*, un entorno de trabajo de código fuente abierto que ofrece una implementación de los protocolos *Liberty* basada en el lenguaje Java.

- **Fase 2, ID-WSF**: En cuanto a la implementación en el prototipo de las funcionalidades de *Liberty* Fase 2, hay que señalar que el entorno de trabajo *SourceID* no cubre esta parte de las especificaciones de Liberty, por lo que ha sido necesario realizar una implementación propia de los protocolos más relevantes.

- **Fase 3, ID-SIS:** Como servicios de *Liberty* Fase 3, se ha implementado uno de los servicios ya especificado por el consorcio (servicio de

[3] "Open Source Federated Identity Management", http://www.sourceid.org.

atributos del Perfil Personal) y se ha desarrollado uno nuevo con aquellos atributos que pueden parecer interesantes en el contexto de aplicaciones de telefonía móvil (servicio de atributos del Perfil de Usuario Móvil).

4.3 Entidades y protocolos Liberty

La estructuración en fases ha impuesto la elaboración de dos escenarios diferentes (siendo el segundo una ampliación del primero), los cuales se describen a continuación.

En la Figura 5 se recoge el escenario que se ha utilizado en la primera parte del proyecto; se representa un *Círculo de Confianza*, compuesto por un Proveedor de Identidad situado en la Red IP privada del Operador y varios Proveedores de Servicios, situados en Internet o en la misma Red IP privada. La conexión entre el Proveedor de Identidad y los Proveedores de Servicios se realiza a través de un *Firewall* que conecta ambos dominios.

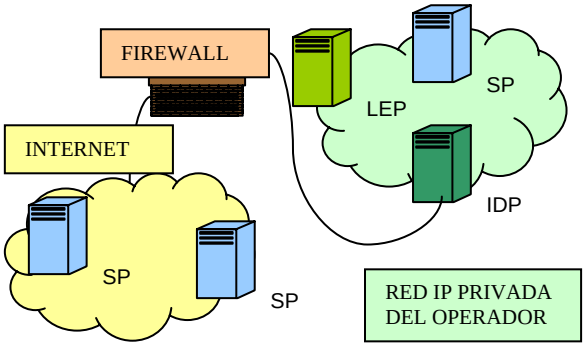


Fig. 5. Escenario de la Fase 1 del proyecto FIRMA.

Por lo tanto, en el contexto de los servicios de telefonía móvil, el *Proveedor de Identidad* es una entidad perteneciente al Operador en la cual el usuario tiene una cuenta establecida y vinculada a sus diferentes cuentas en los *Proveedores de Servicios*, los cuales ofrecen servicios a través de una aplicación ejecutable desde el terminal móvil.

A este diagrama habría que superponerle no obstante una red móvil en la que se sitúan las estaciones base con las que contactan los terminales móviles y una serie de pasarelas entre esta red y la Red IP privada del Operador, como *Gateways*, *MMSC*, *SMSC*, etc.

El *Liberty Enabled Proxy* (LEP) es una entidad situada en el dominio administrativo del Operador de Telefonía Móvil que actúa en nombre del cliente como intermediario en las interacciones que se llevan a cabo entre las diferentes entidades involucradas (Proveedores de Servicios y Proveedor de Identidad).

La utilización de un *Proxy* ofrece dos ventajas fundamentales: la primera es que a la hora de procesar los mensajes *Liberty* (basados en XML) los analizadores de XML trabajan de manera más eficiente en un servidor que en un terminal móvil; y la segunda es que el *Single Sign-On* es un proceso que se lleva a cabo en tres pasos, por lo que serían demasiadas conexiones a establecer desde un terminal móvil, teniendo en cuenta que el modelo de negocio para servicios móviles se suele basar en el volumen del tráfico intercambiado por el cliente.

Los protocolos Liberty utilizados en esta primera fase del proyecto son los ofrecidos por la implementación *SourceID, ID-FF 1.1 Java Toolkit*:

- *Single Sign-On and Federation Protocol [3]*: permite tanto llevar a cabo el establecimiento de sesiones de servicio con Proveedores de Servicio mediante el mecanismo de *Single Sign-On* como establecer la vinculación entre las cuentas del usuario en el Proveedor de Identidad y el Proveedor de Servicio. Se basa normalmente en tres pasos (petición de un mensaje *AuthnRequest* al SP, envío de este mensaje al IDP para obtener un mensaje *AuthnResponse* y envío de este mensaje al SP para finalizar el mecanismo).

- *Single Logout Protocol [3]*: permite cerrar todas las sesiones de servicio iniciadas con el mecanismo de *Single Sign-On* en una sesión de acceso.

- *Federation Termination Protocol [3]*: permite finalizar la vinculación entre las cuentas del usuario en el Proveedor de Identidad y el Proveedor de Servicios.

- *Name Registration Protocol [3]*: permite registrar un nuevo identificador o pseudónimo de la vinculación entre una cuenta en el Proveedor de Identidad y una cuenta en el Proveedor de Servicios.

Para la segunda parte del proyecto hay que superponer algunas entidades adicionales al diagrama anterior, que son las que ofrecen los servicios de autenticación, descubrimiento, acceso a los atributos e interacción con el usuario (ver Fig. 6 y Fig. 7).

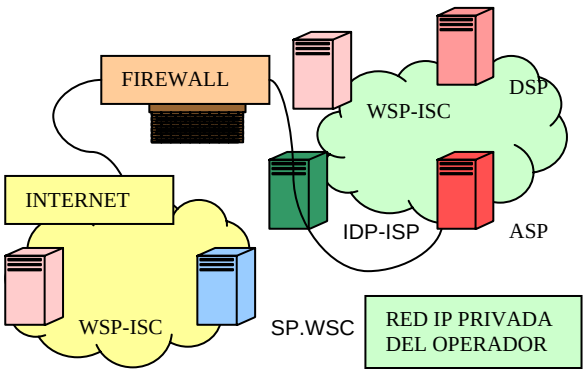


Fig. 6. Escenario de la Fase 2 del proyecto FIRMA.

A continuación se describen las entidades *Liberty* que aparecen en este nuevo escenario.

El *Consumidor de Servicios Web* (WSC) es un Proveedor de Servicios que quiere llevar a cabo una consulta de atributos del usuario o el acceso a un servicio basado en la identidad de éste para ofrecer servicios finales personalizados.

El *Proveedor de Servicios Web* (WSP) es la entidad que ofrecerá al WSC el servicio de consulta de atributos del usuario o llevará a cabo unos servicios basados en la identidad de éste.

El *Proveedor de Servicio de Autenticación* (ASP) es la entidad que permite al WSC autenticarse y conseguir las credenciales necesarias para poder acceder al servicio Web de Descubrimiento.

El *Proveedor de Servicio de Descubrimiento* (DSP) es la entidad que informa al WSC de dónde se encuentra un servicio concreto asociado a un usuario. Gracias al servicio de descubrimiento cada usuario podrá tener los servicios en el WSP que él quiera.

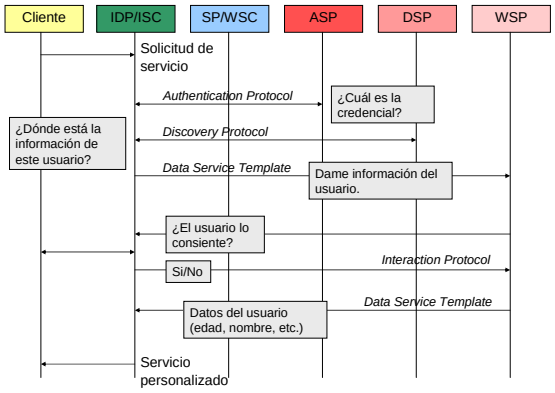


Fig. 7. Diagrama de interacción entre las entidades de la Fase 2.

Los roles de *Proveedor y Consumidor de Servicio de Interacción* (ISP e ISC) aparecen cuando se implementa el protocolo de Interacción. Este protocolo permite a un WSP obtener un permiso de consulta de un usuario para que un WSC pueda acceder a los atributos de éste en caso de que no tuviese ninguna directriz de privacidad anterior. En este sentido, el Proveedor de Identidad desempeñará el papel de ISP ya que es él quien lleva a cabo el diálogo con el usuario y el que remitirá una respuesta al WSP con los nuevos permisos de consulta.

Los protocolos especificados por el proyecto Liberty Alliance de los que hacen uso estas entidades en el marco de trabajo ID-WSF son:

- *Authentication Protocol [4]*: es un protocolo basado en mensajes petición-respuesta que permite implementar diferentes mecanismos de autenticación. Su objetivo es que una entidad proveedora (ASP) pueda dotar de credenciales a un consumidor (WSC) para que pueda acceder a servicios prestados por terceros (DSP). El mecanismo de seguridad implementado en el prototipo es el que se incluye como ejemplo en los documentos de Liberty, el mecanismo CRAM (*Challenge/Response Authentication Mechanism*).

- *Discovery Protocol [4]*: este protocolo es un Servicio Web basado en Identidad desde el punto de vista que permite obtener información sobre un usuario. La información que el proveedor (DSP) ofrece al consumidor (WSC) es información de localización de recursos (WSP); el WSC solicita mediante un mensaje *Query* la localización de un tipo de servicio asociado a un usuario concreto, y el DSP responde con un mensaje *QueryResponse* incluyendo dicha localización y la credencial necesaria para acceder a dicho recurso.

- *Data Service Template [4]*: este protocolo permite llevar a cabo el acceso a los atributos de usuario definidos en la Fase 3 de *Liberty*. Es un protocolo de petición-respuesta (mensajes *Query* y *QueryResponse*) muy simple que permite la interacción entre el WSC y el WSP.

- *Interaction Protocol [4]*: este protocolo permite que un WSP pueda solicitar una interacción con el usuario para recibir nuevos permisos de consulta si no se tenía establecida con anterioridad una directriz de privacidad. En este protocolo el WSP envía un mensaje de petición de interacción al ISP (en esta implementación a través del cliente). Después de llevar a cabo la interacción con el usuario el mismo ISP enviará un mensaje de respuesta con los nuevos permisos al WSP.

Hay que señalar que todos estos protocolos hacen uso de un recubrimiento o envoltorio (*binding*) del protocolo SOAP (*Simple Object Access Protocol) [5]*, para el cual se definen unas cabeceras de *correlación* y de *proveedor*.

4.4 Herramientas y entornos de trabajo

Para el desarrollo del proyecto FIRMA se han utilizado varias herramientas y tecnologías, unas de carácter específico del entorno del proyecto y otras de carácter más general.

Entre las herramientas de carácter específico hay que destacar dos, una empleada en el lado servidor de las aplicaciones y otra en el lado cliente:

- **SourceID, ID-FF 1.1 Java Toolkit.** Este entorno de código fuente abierto ofrece una implementación conforme con las especificaciones de los protocolos definidos por *Liberty Alliance* Fase 1. Consta de una serie de bibliotecas Java que se añadirían a los servidores del Proveedor de Identidad y de los Proveedores de Servicios fácilmente configurables mediante ficheros XML. Estas bibliotecas utilizan

tecnología *Servlet* y *JSP*, lo que permite el acceso a los protocolos *Liberty* tanto desde la ejecución de aplicaciones como desde navegadores de propósito general.

- **KSOAP.** Como implementación de Servicios Web se ha utilizado en este proyecto la ofrecida por las bibliotecas KSOAP [4], que consisten en una serie de clases Java que se pueden utilizar tanto desde el lado Cliente de la aplicación (J2ME) como desde el lado Servidor (*Servlets*). El protocolo SOAP permite de manera sencilla el intercambio de mensajes para el acceso a diferentes operaciones y servicios de una aplicación.

Entre las herramientas o entornos de carácter general hay que señalar la siguientes: para el desarrollo de las aplicaciones en el lado cliente se ha utilizado *Java 2 Micro Edition* con su emulador *Wireless Toolkit*; para el lado servidor de las aplicaciones se ha utilizado el entorno de desarrollo Java *ECLIPSE* con su *plugin* para el contenedor de *Servlets Tomcat* y las bibliotecas JDBC para acceso a base de datos *mySQL* en el lado servidor.

4.5 Demostrador del prototipo

Como demostrador del prototipo se han implementado dos servicios sencillos de Internet móvil que hacen uso de las funcionalidades Liberty. Estos dos servicios consisten en una tienda electrónica y una central de descargas.

Ambos servicios permiten el acceso mediante el mecanismo de *Single Sign-On* ofrecido por un Proveedor de Identidad común que forma parte del Operador de Telefonía Móvil.

El servicio de tienda permite (una vez que el usuario ha accedido a su cuenta) recibir información sobre los productos disponibles (libros, discos, etc.) y solicitar su envío contra reembolso a su domicilio, el cual se consulta de su Perfil Personal haciendo uso de las funcionalidades *Liberty* de Fase 2. Una de las opciones personalizadas de este Proveedor de Servicios es recibir una lista de artículos recomendados en función de la edad del usuario, la cual se consulta nuevamente del Perfil Personal.

El servicio de descargas permite simular la descarga de tonos, logos, canciones y vídeos al móvil teniendo en cuenta los atributos recogidos en el Perfil de Usuario Móvil. Así, cuando un usuario solicita los elementos de descarga disponibles se consulta de este perfil si su terminal móvil tiene capacidad para reproducir MP3, MPEG, qué resolución de color tiene su pantalla, etc. Y en función de esto se le ofrecen unos elementos u otros. Cuando se selecciona para descarga uno de los elementos se consulta del Perfil de Usuario Móvil el MS-ISDN del usuario para poder enviarle la descarga solicitada como MMS.

Para poder gestionar de manera sencilla los atributos de usuario se ha añadido al escenario demostrador un tercer Proveedor de Servicio que permite la modificación de los atributos de usuario, tanto del Perfil Personal como del Perfil de Usuario Móvil. El acceso a esta aplicación también se lleva a cabo mediante el mecanismo de *Single Sign-On*.

4.6 Integración con un servicio ya desarrollado

Como se ha señalado anteriormente uno de los objetivos del proyecto era la validación del prototipo desarrollado mediante la integración de éste con un servicio completo ya desarrollado.

El servicio escogido ha sido un servicio de Mensajería instantánea, basado en tecnologías P2P y de Localización, que permite a los usuarios del servicio enviar y recibir mensajes de su lista de contactos así como obtener la posición en la que se encuentran. Permite también establecer subgrupos de la lista de contactos con los que dialogar y crear grupos de conversación *ad hoc* a los que los usuarios se suscriben por intereses temáticos o por proximidad geográfica.

La integración con este servicio se ha realizado en dos fases:

- **Fase de Integración 1**. Integración de las funcionalidades *Liberty* Fase 1. Esta fase ha consistido en añadir las bibliotecas y los *Servlets* desarrollados en el proyecto FIRMA para el Proveedor de Servicios y las tablas de la base de datos SQL y configurar la aplicación mediante la modificación de los ficheros XML de *SourceID*. También se ha modificado una pequeña parte del código para hacer compatible el *Single Sign-On* con el modelo de sesión de servicio que utilizaba anteriormente el servicio de mensajería.

- **Fase de Integración 2.** Integración de las funcionalidades *Liberty* Fases 2 y 3. Se han añadido las bibliotecas que permiten a un Proveedor de Servicios actuar como Consumidor de Servicios Web y llevar a cabo consultas a un Proveedor de Servicios Web. Esta fase de la interacción conlleva la modificación adicional del código de la aplicación una vez que se han obtenido los atributos, no sólo por la implementación del servicio de interacción, sino porque la consulta de atributos enriquece la oferta de los servicios ofrecidos.

En cuanto a las mejoras que se han añadido al servicio, gracias a la consulta de atributos del usuario

[4] "The home of KSOAP at Enhyndra.org", http://ksoap.objectweb.org/.

con el Perfil de Usuario Móvil se ha conseguido que la aplicación cambie la presentación según el tamaño y la gama de colores soportada por la pantalla del terminal móvil utilizado.

La consulta de los atributos del Perfil Personal permite que los mensajes y el texto de la aplicación presentados al usuario estén en el idioma que éste tuviese seleccionado por defecto. También permite obtener una dirección Web en la cual encontrar una imagen del usuario con el que se está chateando y restringir el acceso a los grupos de conversación según la edad del usuario (se ha tenido que añadir la posibilidad de definir una edad mínima al crear un grupo de conversación).

4.7 Logros y resultados

Como resultado del desarrollo del prototipo, se ha obtenido una validación satisfactoria del entorno de trabajo publicado por *SourceID* para la gestión de identidades en cuanto a las funcionalidades *Liberty* Fase 1 se refiere, tanto por su sencillo acoplamiento con tecnologías *Web Services* como por su facilidad de uso. Así mismo se han obtenido una serie de bibliotecas Java construidas alrededor de este entorno que permiten añadir dichas funcionalidades a una aplicación basada en *Web Services*. Además se ha obtenido una extensión a las bibliotecas de *SourceID* que permiten incorporar las funcionalidades de Fases 2 y 3 de *Liberty*.

Hay que añadir que entre todas estas bibliotecas hay algunas para el lado cliente que proponen una extensión del modelo de desarrollo propuesto por J2ME para la creación de aplicaciones basadas en *Liberty*.

5 Conclusiones y trabajo futuro

Los nuevos servicios sobre redes convergentes plantean nuevos retos técnicos y de negocio. Uno de los aspectos clave en este sentido es el de los nuevos modelos de negocio que aparecen motivados por la creciente competencia y, en el caso de las redes celulares, por el estancamiento en el crecimiento de la tasa de penetración y en la facturación media por abonado. Un referente relevante en este sentido es el modelo de negocio ensayado con éxito en el servicio japonés *i-mode*. En este artículo se presentan algunos resultados interesantes de una línea de investigación que explora la aplicación del modelo de gestión de identidades propuesto por la *Alianza Liberty* para soportar el modelo de negocio *i-mode* y enriquecerlo con una tecnología que permita soportar un modelo de gestión de sesiones en servicios móviles de nueva generación y mejore la experiencia de usuario en un entorno de confianza y seguro.

Con el proyecto FIRMA se ha obtenido una validación positiva del uso de la tecnología *Liberty* con este propósito, pues se ha podido observar a lo largo del proyecto las múltiples semejanzas entre el modelo de Círculo de Confianza y el modelo *semi-walled garden* que tanto éxito ha tenido en aplicaciones de telefonía móvil. La integración con un servicio de Internet móvil ya desarrollado ha demostrado que las funcionalidades *Liberty* no sólo ofrecen comodidad al usuario a la hora de realizar el acceso a los Proveedores de Servicios, sino que permiten ofrecer servicios personalizados de mayor calidad de manera sencilla para el desarrollador de aplicaciones.

El trabajo presentado en este artículo se enmarca en una línea de investigación sobre nuevos modelos de sesión y negocio en servicios avanzados de Internet móvil. Esta línea de investigación está soportada actualmente por la Célula de Innovación @Liberty, proyecto financiado por Ericsson, que plantea el diseño de una infraestructura para la gestión de la federación entre Círculos de Confianza y la gestión de sesiones basada en la tecnología Liberty y los *Web Services* en servicio móviles de nueva generación.

Agradecimientos

Los autores de este artículo desean expresar su agradecimiento a la Cátedra Ericsson de la E.T.S.I. de Telecomunicación de la U.P.M. por su apoyo y financiación a la línea de investigación

Referencias

[1] G. Baker, V Megler, The semi-walled-garden: Japan's "i-mode phenomenon". IBM pSeries Solutions Development, Octubre, 2001.

[2] J. M. Walker Pina, M. A. Monjas Llorente. "Gestión de Sesiones e Identificadores de Usuarios como Mecanismos Centrales en la Red de Servicios de Operadores Móviles.". XIV Telecom I+D, Madrid 23-25 Noviembre (2004).

[3] S. Cantor, J. Kemp. "Liberty ID-FF Protocols and Schema Specification" Version 1.2-errata-v2.0 Liberty Alliance Project, http://www.projectliberty.org/specs.

[4] J. Touzan, Y. Koga. "Liberty ID-WSF Web Services Framework Overview". Version: v1.0-errata-v1.0. Liberty Alliance Project, http://www.projectliberty.org/specs.

[5] J. Hodges, J. Kemp, R. Aarts, "Liberty ID-WSF SOAP Binding Specification" Version 1.1 Liberty Alliance Project, http://www.projectliberty.org/specs.

Uso de resguardos de voto en Sistemas de Votación por Internet

Iñaki Goirizelaia, Iñigo Echave, Maider Huarte, Eduardo Jacob, Juanjo Unzilla
School of Engineering of Bilbao, University of the Basque Country
Alda. Urquijo s/n 48013 Bilbao, Bizkaia
{inaki.goirizelaia, jtpecgui, maider.huarte, eduardo.jacob, juanjo.unzilla@bi.ehu.es}
Tfno:+34946014210

Technology is not mature enough to accept Remote Internet Voting anywhere anytime. In such situation there is no way to offer voters a coercion free voting system. On the other hand, Internet Voting Systems may help improving efficiency of electoral processes while offering electors mobility to cast their votes from anywhere. But if that is the case, how do we make sure that our vote is really counted? The use of Vote Receipt Trails is proposed to answer this question. This paper presents an Internet Voting System that in order to avoid coercion requires the use of voting kiosks controlled by electoral authorities and it also offers a vote receipt that can be used to check that every vote is counted but it cannot be used to buy or sell votes.

1 Introducción

La tradición electoral en la mayoría de los estados, se basa en el uso de votos de papel, pero hay buenos ejemplos de estados democráticos en los que el proceso electoral se está cambiando del papel a los votos digitales. Recientemente, hemos sabido que un referéndum en Venezuela (2004) se hizo usando sólo votos electrónicos. Durante las últimas elecciones presidenciales de los EEUU (2004), se estimó que un 63.7% de los electores usaron máquinas electrónicas para dar su voto, de las cuales, un 30% eran sistemas OMR (*Optical Mark Recognition*, Reconocimiento Óptico de Marcas) y el 33.7% restante, máquinas basadas en pantallas táctiles [1]. La tecnología en los procesos electorales se está generalizando en un gran número de estados democráticos aunque algunas organizaciones digan que el uso de máquinas de voto electrónico es la mejor y más fácil forma de atacar a la democracia.

Los Sistemas Electrónicos de Votación están abriéndose paso en diferentes estados pero la gente necesita más tiempo para aceptar cambios tan importantes en la manera de votar. Hoy en día, nadie duda de que para mejorar los procesos electorales, la tecnología es necesaria. La pregunta es, ¿qué clase de tecnología? Este artículo presenta un Sistema de Votación por Internet que intenta reproducir la forma tradicional de votar. El informe *Report on the Feasibility of Internet Voting* [8] fue usado como guía en nuestro diseño. Este informe propone una arquitectura segura para el Voto por Internet y recomienda introducirlo de forma gradual. Presenta diferentes maneras de Voto por Internet en el lugar donde esté el votante (en cualquier lugar, en Quioscos electorales precintados o controlados en cualquier sitio). Esta propuesta presenta un Sistema de Votación por Internet donde los electores pueden dar su voto desde cualquier Quiosco en cualquier colegio electoral. Este significa que los votantes pueden elegir el lugar desde donde quieren emitir su voto, obteniendo flexibilidad y movilidad.

Cualquier Sistema de Voto por Internet está basado en un protocolo de voto específico que debería ser diseñado para mantener la privacidad del votante y evitar el ataque de cualquier hacker. Algunos protocolos de voto desarrollados hasta el momento se basan en el uso de canales anónimos de votación. Estos protocolos están diseñados de forma que es imposible relacionar un voto con el votante que lo emitió. Este tipo de protocolos requieren por lo menos dos fases. En la primera, los votantes reciben una especie de testigo firmado ciegamente por los administradores de las elecciones. Durante la segunda fase, ese testigo se usa para autenticar el voto encriptado transmitido a los agentes contadores a través del canal anónimo. De hecho, muchos protocolos de votación propuestos se basan en este esquema [2, 3, 4, 5].

Hay otros protocolos de votación donde los electores tienen que identificarse a sí mismos pero el proceso de contado, que es universalmente verificable, asegura que no hay forma de relacionar la identidad del votante con el contenido de su voto. En ese caso, emitir el voto significa enviar un simple mensaje a un panel electrónico. Para conseguir privacidad y secreto de voto, se proponen el uso de encriptación homomórfica [6, 7, 14] y redes de mezclado [9].

Prevenir la coacción (“espiar por encima del hombro”, coacción por proximidad, voto familiar, etc.) es clave para el éxito del Voto por Internet. La compra del voto está siempre relacionada con el concepto de no-resguardo. Los intimidadores quieren saber que un votante emite su voto de la forma que ellos quieren. Si se necesita una cabina para emitir el voto, esta forma de coacción se puede evitar simplemente usando esquemas sin resguardo. Si fuéramos capaces de ofrecer sistemas que no provean ninguna prueba de cuál fue la opción, no hay forma

de vender el voto. El concepto de no-resguardo fue presentado por primera vez por Benaloh y Tuinstra [10], basándose en el uso de una cabina de votación y un protocolo de votación con encriptación homomórfica. Se han propuesto muchas soluciones basadas en el concepto del no-resguardo; la mayoría de ellas están basadas en dos asunciones: un canal inviolable desde el votante hasta la autoridad electoral y viceversa y/o una cabina de votación donde los votantes emiten su voto [12, 13, 14]. Obviamente, si los votantes usan cualquier tipo de dispositivo conectado a Internet para emitir su voto, no hay forma de prevenir que un intimidador esté mirando a los votantes. Hoy en día no hay solución técnica para ese problema.

Chaum presentó un nuevo tipo de resguardo [11]. Su propuesta se basa en el uso de un resguardo de papel de dos capas, que cuando se presentan juntas indican la opción seleccionada por el votante pero que al separarse, no hay forma posible de leer dicha opción. El votante ha de separar las capas al salir de la cabina y una de ellas se la dará al encargado del colegio, para que la destruya. Una versión electrónica del último resguardo exacto, que codifica el voto, se mantiene en la máquina de votación. Dentro de la cabina de votación, el resguardo de dos capas es tan convincente como cualquier resguardo. Fuera de la cabina, el resguardo está constituido por una única capa, y sólo puede ser usado para comprobar que el voto emitido fue incluido correctamente en la cuenta final. Pero no puede ser usado para demostrar el contenido del voto emitido al intimidador.

Las coacciones podrían neutralizarse usando smart-cards y algún tipo de PIN de alerta. En ese caso, el votante podría emitir su voto usando dos PINs diferentes y el sistema de votación sólo acepta votos emitidos con el PIN correcto. Al seleccionar el PIN de alerta, todo parecerá normal, y el votante podrá enseñar al intimidador que está votando como él quiere, pero ese voto jamás será contado.

El uso de sistemas biométricos también ha sido propuesto como forma de prevenir la coacción en los Sistemas de Votación remotos. Parámetros como la voz (tono, análisis espectral) o la firma (presión, velocidad) pueden ayudar para ver si el votante está teniendo un comportamiento inusual debido a una coacción. Sin embargo, el uso de la biometría está muy lejos de ser una opción real en un futuro cercano. Los Niveles de Rechazos Falsos (rechazos que no deberían ser tales) y Niveles de Aceptaciones Falsas (aceptaciones que deberían haber sido rechazadas) son todavía demasiado altos como para ser aceptados en los Sistemas de Votación.

Nosotros creemos que la tecnología no es lo suficientemente madura como para aceptar el Voto Remoto por Internet en cualquier lugar y en cualquier momento. En una situación así, no hay forma de ofrecer a los votantes un Sistema de Votación sin coacción. Si embargo, los Sistemas de Votación por Internet pueden ayudar a mejorar la eficiencia de los procesos electorales, a la vez que ofrecen movilidad a los electores para poder emitir su voto desde cualquier lugar. Pero en ese caso, ¿cómo estamos seguros de que nuestro voto ha sido realmente contado? Nosotros proponemos el uso de un tipo especial de resguardos de voto para responder a esa pregunta. Este artículo presenta un Sistema de Votación por Internet que, para evitar la coacción, requiere el uso de Quioscos de Votación controlados por autoridades electorales y también ofrece un resguardo de votación que puede usarse para comprobar que todos los votos fueron contados pero que no puede usarse para vender o comprar votos.

2 Arquitectura y protocolo de votación del Bolanta Remote Voting System

Nuestra propuesta se basa en el uso de firmas ciegas, sobre todo porque, usando esta técnica, el procedimiento que el votante deberá seguir para emitir su voto, será más fácil de aceptar. La arquitectura del Sistema de Votación Bolanta es similar al propuesto por Fujioka, Okamoto y Otha [3], o el propuesto por Lorrie Cranor [2] pero introduce muchas mejoras que se explicarán a continuación.

2.1 Arquitectura del Sistema de Votación

Los agentes principales de la arquitectura de nuestro Sistema de Votación son los siguientes (ver Fig. 1):

Quiosco de Votación
Los Quioscos de Votación se sitúan en las oficinas electorales y cualquier votante puede utilizarlos. El número de Quioscos de Votación es ilimitado.

Agente de Registro
Este agente se asegura de que el votante es un votante registrado y que no ha votado hasta el momento. Una vez que comprueba que esta persona puede votar, le da el permiso necesario para interactuar con el interfaz de votante. El sistema ha sido diseñado para aceptar más de un Agente de Registro y en ese caso, cada uno de ellos debe dar su permiso al votante. Esto mejora la seguridad y confianza en el sistema. El votante necesitará conseguir, el permiso de, al menos, la mitad de los Agentes de Registro, para poder emitir su voto.

Urna Digital Virtual
Las urnas Digitales Virtuales son el equivalente a las urnas tradicionales. Almacenan todos los votos digitales emitidos por los votantes, antes de que sean tabulados. Es importante señalar que cada Urna Virtual sólo almacena los votos asignados a sí misma. Así, se reproduce la forma en la que se organizan las elecciones hoy en día. También proporcionan un resguardo de votación al votante, que éste puede usar,

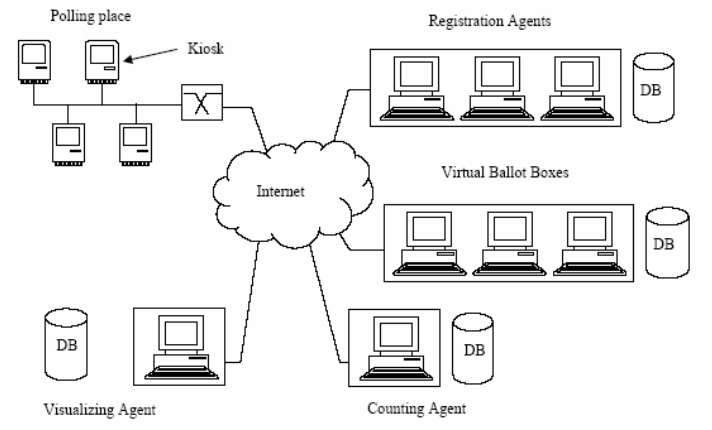


Fig. 1: Arquitectura del Sistema de Votación Bolanta

una vez que se hayan publicado los resultados electorales, para comprobar si su voto fue contado. El concepto de Distrito Electoral está asociado a cada Urna Virtual, posibilitando así un esquema sencillo para contar los votos, de la misma forma que se hace tradicionalmente, usando urnas de votos de papel. Se pueden usar tantas Urnas Digitales Virtuales como se necesiten y cada una de ellas usa el mismo certificado digital.

Agente de Contado
Este agente cuenta todos los votos de cada Urna Virtual. Primero, recibe todos los votos y resguardos de cada Urna Virtual. Después, tabula los resultados para cada Urna Virtual y los suma de forma jerarquizada, calculando así resultados parciales y totales. Finalmente, transmite todos los resultados y resguardos de voto al Agente de Visualización.

Agente de Visualización
Este agente crea todas las páginas web dinámicas que mostrarán los resultados de la elección. También facilita al votante las herramientas necesarias para comprobar si su voto fue contado o no.

2.2 Protocolo de Votación

El proceso de votación comienza cuando el votante llega al Quiosco de Votación. A continuación, se indica una descripción del protocolo de votación y el proceso seguido por el votante (ver Fig. 2):

1. El votante teclea su código secreto, que es único y ningún otro votante puede usar. Este proceso podría mejorarse usando DNI-s digitales o smart-cards.

2. El Quiosco de Votación, usando el código secreto del votante, intenta obtener el permiso necesario para seguir con el proceso de votación. Genera un número aleatorio, lo firma digitalmente, lo ciega y lo transmite a todos los Agentes de Registro del sistema, junto con el código del votante. El número aleatorio firmado y cegado es la petición de permiso:

   code, $sign_K(blind(sign_K(R)))$

   Donde $sign_K$ indica que el Quiosco (K) lo firmó.

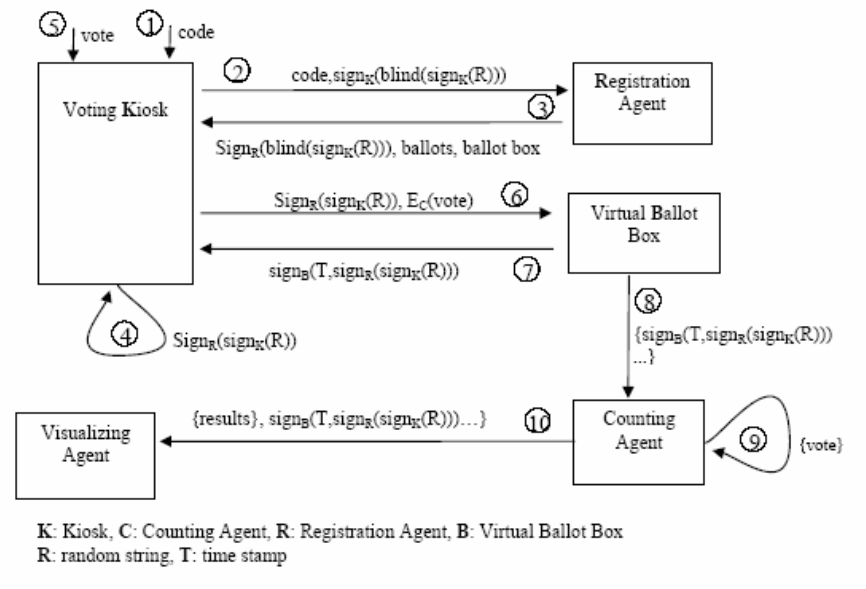


Fig. 2: Descripción del Protocolo de Votación

3. Todos los Agentes de Registro, siguiendo procesos independientes, obtienen la identidad del votante basada en su código y comprueban si tiene derecho a votar, es decir, si es un votante registrado que no ha votado aún. Si ese es el caso, firman de forma ciega la petición de permiso y se la envían al Quiosco de Votación, junto con las papeletas de votación digitales y la información sobre la Urna Virtual donde se depositará el voto del votante.

   $sign_R(blind(sign_K(R)))$, *ballots, ballot box info*

   Donde *ballots* son las papeletas digitales y *ballot box info* la información sobre la Urna Virtual.

4. Cuando el Quiosco de Votación recibe las respuestas de todos los Agentes de Registro, comprueba que las firmas de los mismos son válidas y que el número de permisos obtenidos es mayor que la mitad del número de Agentes de Registro. Por otra parte, también comprueba que la información sobre la Urna Virtual y las papeletas digitales asignadas al votante por cada uno de los Agentes de Registro, son las mismas en todos los casos. Finalmente, descubre (desciega) la petición de permiso y obtiene todos los permisos firmados por los Agentes de Registro.

   $sign_R(blind(sign_K(R))) \rightarrow sign_R(sign_K(R))$

5. El Quiosco de Votación presenta todas las papeletas de votación al votante. El acto de votar consiste en elegir una de las papeletas, la que pertenece al partido político al que el votante quiere votar.

6. Después, el voto es encriptado usando la clave pública del Agente de Contado y se transmite a la Urna Virtual asignada al votante según su distrito electoral.

   $sign_R(sign_K(R))$, *EC(vote)*

   Donde ***EC*** significa que el contenido del voto (*vote*) está encriptado usando la clave pública del Agente de Contado.

7. Cuando la Urna Virtual recibe un voto, comprueba que todos los permisos de votación del mismo son válidos y el número de ellos es suficiente. Si ese es el caso, la Urna Virtual genera un resguardo de voto basándose en los permisos, un sello de tiempo y su firma digital. Así, es imposible conseguir un resguardo de voto sin haber emitido el voto desde un Quiosco, sin el número suficiente de permisos de los Agentes de Registro y sin haber mandado el voto a la Urna Virtual asignada al votante.

   $sign_B(T, sign_R(sign_K(R)))$

8. Cuando el día electoral acabe, cada administrador electoral se autenticará ante el agente que estuviera bajo su control. Cuando un agente detecta que todos los administradores electorales asignados a él se han autenticado correctamente, deja de ofrecer su servicio. Desde ese momento, las Urnas Virtuales no aceptarán más votos y transmitirán todos los votos y resguardos almacenados en su Base de Datos al Agente de Contado.

9. El Agente de Contado desencripta todos los votos usando su clave privada. Después, este agente cuenta los votos, teniendo en cuenta que cada voto pertenece a un distrito electoral concreto y que ha de calcular cuentas parciales. Una vez realizadas todas las cuentas (parciales y totales) se transmiten junto con los resguardos al Agente de Visualización.

10. El Agente de Visualización recibe todos los resultados y resguardos de voto del Agente de Contado, y los almacena en su Base de Datos.

11. Cuando un votante quiere ver los resultados de las elecciones, puede hacerlo visitando la página web publicada por el Agente de Visualización. También es posible comprobar si el resguardo de voto del votante está almacenado en la base de datos del Agente de Visualización. Si un votante puede ver su resguardo de voto en el Agente de Visualización, eso significa que su voto ha sido contado, sin que nadie haya podido robarlo.

3 Resguardos de Voto

El Sistema de Votación por Internet Bolanta, propone una nueva forma de expedir resguardos de voto (ver Fig. 3) que ha sido diseñada para garantizar lo siguiente:

- Todos los votos fueron emitidos desde Quioscos legales.
- El votante tiene el permiso de la mayoría de los Agentes de Registro.
- El voto es almacenado en la Urna Virtual asignada al votante.
- Los Agentes de Registro no pueden emitir votos usando las identidades de los votantes que decidieron abstenerse, porque no conocen la clave privada de ningún Quiosco, que se usa para firmar el permiso.
- El resguardo no tiene información sobre el contenido del voto.

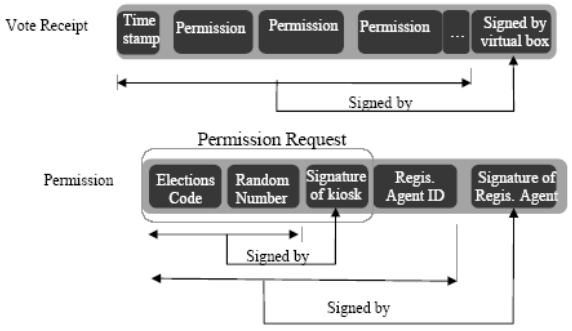


Fig. 3: Estructura de un Resguardo de Voto y cada uno de los permisos que la componen

- El resguardo no puede ser usado en otras elecciones o en las mismas, otra vez.

El resguardo se compone de los permisos obtenidos desde los Agentes de Registro junto con un sello de tiempo y un código correspondiente a las elecciones. Esos permisos, sello y código son firmados por la Urna Virtual asignada al votante. A su vez, cada permiso dentro del resguardo está firmado por el Quiosco desde el que se pidió y el Agente de Registro que lo concedió.

4 Conclusiones

Este artículo presenta un Sistema de Votación Remoto que usa Quioscos controlados por administradores locales para la emisión del voto. El sistema proporciona un resguardo de voto que puede ser usado por el votante para comprobar que su voto fue contado. Los votantes pueden emitir su voto desde cualquier Quiosco en cualquier colegio electoral.

La arquitectura general de nuestro sistema es similar al propuesto por Fujioka, Okamoto y Otha o Lorrie Cranor (conocido como Sensus), ambos basados en el uso de firmas digitales, pero incluye una serie de importantes contribuciones, indicadas a continuación:

- Cada votante tiene un código privado único que necesita teclear en el Quiosco para autenticarse ante la mayoría de los Agentes de Registro, antes de poder emitir su voto. Una vez que el votante es autenticado, se le informa de cuál es el distrito electoral donde está registrado.

- Se propone un nuevo agente, la Urna Virtual, que comprueba que cada votante tiene los permisos necesarios de los Agentes de Registro para votar, y también genera los resguardos de voto y almacena todos los votos correspondientes a un distrito electoral. En comparación con una urna electoral tradicional, la Urna Virtual ofrece la posibilidad de recibir votos desde cualquier Quiosco localizado en cualquier colegio electoral, proporcionando una flexibilidad que el sistema tradicional no tiene.

- Después de emitir el voto, la Urna Virtual proporciona al votante su resguardo de voto. La forma en la que este resguardo de voto se genera, es una de las contribuciones de este trabajo. Usando este resguardo de voto, no hay forma de saber el contenido del voto, no se puede volver a usar ni se puede falsificar. La seguridad y la fiabilidad del sistema es también mejorada, ya que gracias a estos resguardos de voto, es imposible que los Agentes de Registro cometan fraude votando por las personas que decidieron no hacerlo.

Agradecimientos

Los autores quieren agradecer a la Dirección General de Procesos Electorales del Gobierno Vasco por el apoyo y patrocinio dado a este trabajo.

Referencias

[1] "Overview of Voting Equipment Usage in United States". Election Data Services. http://www.electiondataservices.com/EDSInc_DREoverview.pdf, Mayo 2004.

[2] Sensus. http://lorrie.cranor.org/voting

[3] A. Fujioka, T. Okamoto, K. Ota, A practicalSecret Voting Scheme for Large Scale Election, *Advances in Cryptology-AUSCRYPT'92*, Lecture Notes in Computer Science vol.718, p.248-259,Springer-Verlag, 1993.

[4] Kwangjo Kim, Jinho Kim, Byoungcheon Lee, Gookwhan Ahn. "Experimental Design of Worldwide Internet Voting System using PKI", SSGRR2001, L'Aquila, Italia, Agos. 6-10, 2001

[5] EVOX. http://theory.lcs.mit.edu/~cis/voting/protocol/index.htm

[6] Adler Jim, Dai W, Green R.L., Neff C.A.;"Computational Details of the VoteHere Homomorphic Election". http://votehere.net November 2000.

[7] ElGamal, T.; "A public-key cryptosystem and a signature scheme based on discrete logarithms." IEEE Transactions on Information Theory, IT-31 (4):469-472, 1985.

[8] California Internet Voting Task Force, "Report on the Feasibility of Internet Voting", http://www.ss.ca.gov/executive/ivote/, 2000.

[9] Jakobsson M., Juels A., Rivest R.L.; "Making mix nets robust for electronic voting by randomized partial checking". Proceedings of the 11th USENIX Security Symposium, ISBN 1-931971-00-5, pp 339-353, 2002

[10] Benaloh, J.C., Tuinstra D.; "Receipt-free secret ballot elections". Proceedings 26th ACM Symposium on the Theory of Computing (STCO), pp 544-553. ACM 1994.

[11] Chaum D.; "Secret-Ballot Receipts and Transparent Integrity". IEE Security & Privacy. Enero-Febrero 2004, Vol 2 N1, pp 38-47.

[12] Magkos E., Burmester M., Chrissikopoulos V., "Receipt-freeness in Large-scale elections without untappable channels". Proc. 1st. IFIP Conference on E-Commerce / E-Business / EGovernment, pp 683-693. Kluwer Academics Publishers, 2001.

[13] Ku WC., HO CM., "An e-Voting Scheme against Bribe and Coercion". Proceedings of the 2004 IEEE International Conference on e-Technology, e-Commerce and e-Service. 2004.

[14] Hirt M., Sako K.; "Efficient Receipt-Free Voting Based on Homomorphic Encryption". EUROCRYPT 2000, vol 1807 of LNCS, pp. 539-556. 2000

Efficient Certificate Revocation System Implementation : Huffman Merkle Hash Tree (HuffMHT)

Jordi Forné Jose L. Muñoz Manel Rey Oscar Esparza
Technical University of Catalonia (UPC)
Telematics Engineering Department (ENTEL)
1-3 Jordi Girona, C3 08034 Barcelona (Spain)
Phone. +34934010804 Fax.+34934011058
{jose.munoz, jordi.forne, oscar.esparza}@entel.upc.es

Abstract *The public-key is usually made public by way of a digital document called Identity Certificate (IC). ICs are valid during quite long periods of time. However, there are circumstances under which the validity of an IC must be terminated sooner than assigned and thus, the IC needs to be revoked. In this paper, we present practical aspects of a certificate revocation system called Huffman Merkle Hash Tree (HuffMHT). HuffMHT provides an efficient and balanced performance with regards other proposals in the sense that the system does not save bandwidth at the expense of processing capacity and viceversa. Finally, some performance results of HuffMHT are exposed as well.*

1 Introduction

A Public Key Infrastructure (PKI) is required to securely deliver public-keys to widely-distributed users or systems. The public key is usually made public by way of a digital document called Identity Certificate (IC). The PKI is responsible for the Identity Certificates (ICs) not only at the issuing time but also during the whole life-time of the certificate. An IC has a bounded life-time: it is not valid prior to the activation date and it is not valid beyond the expiration date. Typically, the validity period of an IC ranges from several months to several years. In this context, certificate revocation can be defined as "*the mechanism under which an issuer can invalidate the binding between an identity and a public-key before the expiration of the corresponding certificate*". Thus, the existence of a certificate is a necessary but not sufficient evidence for its validity, the PKI needs to provide its end users with the ability to check, at the time of usage, that certificates are still valid (not revoked). This feature is commonly known in the PKI as the status checking.

The Revocation Dictionary ($\mathcal{RD}$) can be defined as the cryptographic structure that contains the status data about the revoked certificates of the PKI domain. The master copy of the $\mathcal{RD}$ for a set of certificates is updated by a Trusted Third Party (TTP) called "issuer". The update process must reflect the revocations and expirations (if a certificate has expired it makes no sense to store revocation information about it). The $\mathcal{RD}$ issuer is also responsible for making publicly available the status data. Usually, the end entities that want to perform a status checking do not have a straight connection to the issuer, they get the status data from intermediate entities instead. In this sense, the issuer can distribute the $\mathcal{RD}$ using two kind of intermediate entities:

- **Repositories (offline status checking).** In this case repositories are not TTPs because the cryptographic evidence for the status data is previously produced by a trusted issuer. The simplest structure for offline distribution is a signed "black" list that includes all the identifiers (serial numbers) of all revoked but not expired certificates issued by the PKI domain. There are several standards based on this idea, below we mention them.

  Traditional Certificate Revocation List (CRL) is the most mature offline system. CRL is part of X.509 [10] and it has also been profiled for the Internet in [1]. A CRL is a digitally signed list of revoked certificates where for each entry within the list the following information is stored: the certificate serial number, the revocation reason and the revocation date.

  Delta-CRL (D-CRL) [3] is an attempt of reducing the size of the CRLs. A Delta-CRL is a small CRL that provides information about the certificates whose status have changed since the issuance of a complete list called Base-CRL.

  In *CRL-Distribution Points* (CRL-DP) [3] each list contains the status information of a certain subgroup of certificates and each subgroup is associated with a distribution point. Each certificate has a pointer to the location of its distribution point.

  The *Certificate Revocation Tree* (CRT) [4]

and the *Authenticated Dictionary* (AD) [9] are both based on hash trees [5]. The hash tree allows content to be retrieved in a trusted fashion with only a small amount of trusted data. The content is stored in the leaves of the hash tree but only the root of the tree is trusted (this structure is further discussed in the next Section).

- **Responders (online status checking).** In this case the cryptographic evidence for the status data is produced online by the responder, that is to say, responders are TTPs. Online schemes usually use the responder's signature over the status data as cryptographic evidence. Notice that end entities are not required to be aware of the back-end infrastructure used to collect the revocation information and maintain the responder's local database[1]. The most popular online protocol used by responders is the *Online Certificate Status Protocol* (OCSP) [8] that has been proposed by the PKIX workgroup of the IETF.

Many benefits can be found in offline status checking. Since repositories are not TTPs, there is not a private key to be protected and the compromise of a repository does not compromise the security of the revocation system. Furthermore, it is easy to add redundancy for status checking in the PKI domain and provides a low-cost status checking in terms of processing capacity necessary in the repository.

Traditionally, the main drawback of offline systems is the communication overhead introduced in the status checking which hinders its development in bandwidth-constrained environments (such as m-commerce). In this paper, we present Huffman Merkle Hash Tree (HuffMHT), a revocation system that provides an efficient status checking using repositories. HuffMHT presents the inherent advantages of offline distribution and also keeps a good performance in terms of processing capacity and communication overhead. Besides, HuffMHT is not constrained by large populations, being scalability one of its main advantages and becoming therefore a realistic and practical system. HuffMHT uses the statistics of the status checking, like in the Huffman algorithm for source coding, for building an unbalanced MHT that minimises the average communication overhead in the status checking process.

The rest of the paper is organised as follows: in Section 2, we describe AD-MHT and HuffMHT. In Section 3, we introduce the most relevant practical aspects that must be taken into account to implement HuffMHT and the design decisions made. In Section 4, we present performance results of HuffMHT compared to AD-MHT that show the better performance of HuffMHT. Finally, we conclude in Section 5.

2 Related Work

2.1 ADMHT

We need first to describe AD-MHT [6], an implementation of a certificate revocation system that uses the data structures proposed by Naor and Nissim in their Authenticated Dictionary (AD) [9]. AD-MHT is based in a balanced hash tree. A sample balanced hash tree is depicted in Figure 1.

We denote by $N_{i,j}$ the nodes within the tree here i and j represent respectively the i-th level and the j-th node. We denote by $H_{i,j}$ the cryptographic value stored by node $N_{i,j}$. Nodes at level 0 are called "leaves" and they represent the data stored in the tree. In the case of revocation, leaves represent the set Φ of certificates that have been revoked: $\Phi = \{s_0, s_1, \ldots, s_j, \ldots, s_{n-1}\}$. Here s_j is the data stored by leaf $N_{0,j}$. Then, $H_{0,j}$ is computed as (1)

$$H_{0,j} = h(s_j)\,. \tag{1}$$

Here h is a One Way Hash Function (OWHF). To build the tree, a set of k adjacent nodes at a given level i; $N_{i,j}$, $N_{i,j+1}$, ...,$N_{i,j+k-1}$, are combined into one node in the upper level, node that we denote by $N_{i+1,j}$. Then, $H_{i+1,j}$ is obtained by applying h to the concatenation of the k cryptographic variables (2)

$$H_{i+1,j} = h(H_{i,j}|\, H_{i,j+1}|\, \ldots |\, H_{i,j+k-1}). \tag{2}$$

At the top level there is only one node called "root". The $\mathcal{D}$igest of the tree is defined as the H_{root} value and a validity period signed by the issuer. The $\mathcal{P}\text{ath}_{s_j}$ can be defined as the set of cryptographic values necessary to compute H_{root} from the leaf s_j.

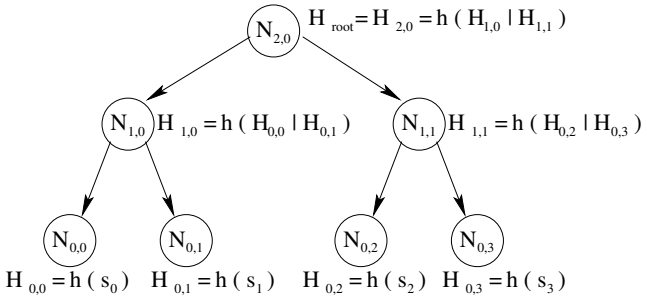
Example. Let us suppose that a certain user wants to find out if s_1 belongs to the sample tree of Figure 1. Then $\mathcal{P}\text{ath}_{s1} = \{H_{0,0}, H_{1,1}\}$ and the response verification consists in checking that the $H_{2,0}$ computed from the $\mathcal{P}\text{ath}_{s_1}$ matches $H_{2,0}$ included in the $\mathcal{D}$igest

$$H_{root} = H_{2,0} = h(h(h(s_1)|H_{0,0})|H_{1,1})\,. \tag{3}$$

Notice that the hash tree can be pre-computed by a TTP and distributed to a repository because a leaf cannot be added or deleted to the $\mathcal{RD}$ without modifying H_{root}[2] which is included in the $\mathcal{D}$igest.

[1] The responder database is usually updated by means of a CRL or requesting other responders.

[2] To do such a thing, an attacker needs to find a pre-image of a OWHF which is by definition computationally infeasible.



Note: h is a OWHF

Figure 1: Sample AD-MHT

The sample tree of Figure 1 is a binary tree because adjacent nodes are combined in pairs to form a node in the next level ($k = 2$).

2.2 HuffMHT

Despite the good behaviour of balanced hash trees compared to CRLs, they have higher communication costs than online systems which is still a problem in bandwidth-constrained environments. HuffMHT uses the statistics of the status checking for building an unbalanced hash tree. The idea is to provide shorter paths for the leaves that have the higher request rates. This structure minimises the average length of the membership response provided by the $\mathcal{RD}$ compared to balanced hash trees [7].

The unbalanced hash tree performs better than the balanced hash trees when the membership of certain elements of the dictionary is verified more frequently than other elements. In the case of revocation this might happen in many scenarios, for instance, in the Business-to-Consumer scenario (B2C) where status data of the servers' certificates is requested more often compared to clients'. Anyway, in the worst case (the request rate is equiprobable for all the data contained by the tree) our approach leads to a binary balanced tree.

Below, we outline the algorithm[3] that builds the hash tree.

Let us assume that Π_i is the probability for membership of element s_i to be requested, then

1. Line up the set of elements by falling probabilities Π_i.
2. The two elements with least probabilities are combined to generate a new node as explained in the previous Section. The new node (a internal tree's node) now is considered to have a probability the sum of probabilities of the two elements.
3. Go to the first step until a single node which probability is 1 is generated. This element will be the root of the tree.

3 Key Implementation Aspects

The basic idea of the HuffMHT has already been exposed, but several implementation decisions have to be properly addressed in order to implement a practical system. In particular, the most critical problems to solve are derived from the fact that tree adjacent leaves do not contain consecutive serial numbers. In this section, we briefly describe several key aspects that we had to face during the implementation phase.

3.1 Content of the tree leaves

The AD-MHT implementation stores single certificates in the tree leaves which are ordered by serial number in the last level. This eases searching information within the tree and supports the dynamism of the tree. To demonstrate the validity of a certain certificate, the next evidences must be given:

1. The existence of a minor adjacent to the target certificate which is revoked.
2. The existence of a major adjacent to the target certificate which is revoked.
3. It must be demonstrated that those previous adjacent leaves are effectively adjacent in the tree.

However, this "adjacency checking" is not practical for the HuffMHT, because the leaves in the last level of the tree, instead of being ordered by their serial number, are randomly ordered depending on their probability. To overcome this problem, HuffMHT stores serial number intervals in the tree leaves, making "adjacency checking" not necessary.

[3]The algorithm we use to build the unbalanced binary tree is equivalent to the one used by Huffman in the binary coding [2].

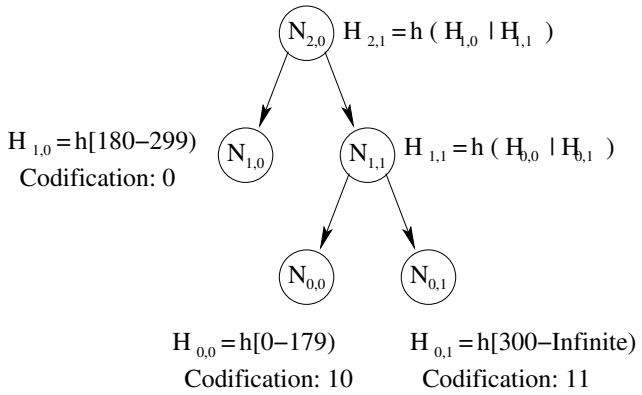


Figure 2: A sample HuffMHT tree. The certificates with serial numbers 180 and 300 are revoked. The interval [180-299) is checked with probability 0,5, while the others are checked with probability 0,25 each one.

Table entry	Minor certificate	Major certificate	Codification
#1	0	179	10
#2	180	299	0
#3	300	Infinite	11

Table 1: Associate list (AL) of the sample HuffMHT tree shown in Figure .

3.2 Searching leaves in the tree

The leaves of the AD-MHT are sorted and, therefore, is quite straightforward to implement a searching algorithm of complexity $O(log(N))$ for a tree with N leaves. However, due to its unpredictable topology, the HuffMHT seems to lack of the necessary dynamism to implement an efficient searching within it. In this section, we will show that the same codification used to build the tree can be easily used to implement an efficient leaf-searching algorithm. This idea leads us to implement a practical revocation system with a computational overhead in the server comparable with the AD-MHT system.

There are two main data structures in the HuffMHT: the hash tree and the associated list. The "associated list" (AL) is a sorted list which stores all the necessary information to efficiently reach a target leaf within the tree. The AL is formed by tuples of three values (4):

$$< minor\,cert,\ major\,cert,\ codification > \quad (4)$$

The codification shows the path to follow from the root to reach the leaf storing the revocation information, the cryptographic value, of that interval. The criterion applied to get the codification of a leaf is: being in a parent node, if codification indicates '1' redirect to the right child and redirect to the left child if codification indicates '0'. Note that the codification of a leaf exactly matches with its associated binary Huffman code.

With respect to the hash tree, its topology depends exclusively on the requests statistics (unbalanced). A sample HuffMHT tree is shown in Figure 2:

It can be seen that leaves can be found in any level and each has its codification associated. Table 1 shows the AL corresponding to this sample tree. The table is sorted, allowing a searching algorithm of complexity $O(log(N))$. When a certificate is found in the AL, its associated codification directly indicates its position in the HuffMHT. The searching algorithm is as simple as starting at the root node and, depending on the bit of the codeword; move to the right (1) or to the left (0), form the first to the last bit of the codeword.

For example, let us assume that we are requesting about the status of the certificate with a serial number of 150. From the AL we obtain that 10 is the codification associated from this range. To search the position of the leaf in the tree, we begin in the root node, then the first 1 lead us to $N_{1,1}$, and finally the last 0 lead us to node $N_{0,0}$.

3.3 Tree set-up and update

Before building the initial tree, we must know the probability of each leaf of the tree. Usually, when the system starts up, these probabilities are not known. In this case, the leaves can be considered equiprobable, leading to a balanced tree. Later, adaptive algorithms can be used to learn the actual probability of each leaf through statistical monitoring (for this purpose, counters in the repositories can be used to inform the issuer).

On the other hand, when a certificate has been revoked or when a revoked certificate reaches the end of its documented life-time, the status data must be updated. The tree is periodically rebuilt to include updated data and the rebuilding process is performed according to the collected statistics.

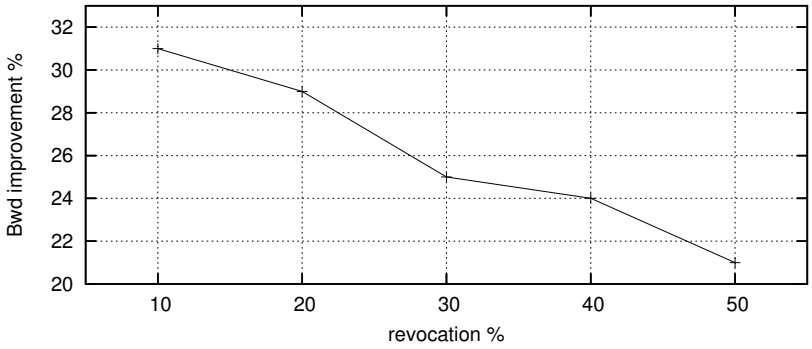


Figure 3: HuffMHT BW_d improvement related to the revocation % of the population.

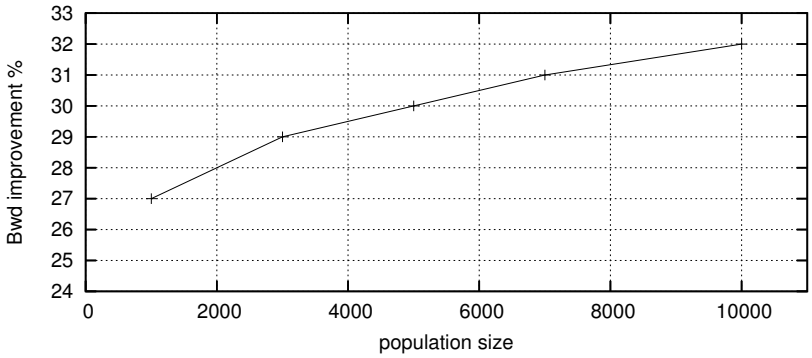


Figure 4: Scalability performance.

4 Evaluation

In this section we compare HuffMHT with the AD-MHT. Results are shown as the % of the difference between both systems (always HuffMHT performing better). BW_d refers to the bandwidth necessary in the server downlink for a given situation.

In the evaluation process we have isolated the savings in BW_d due to the different design decisions. In this first set of simulations, the Huffman encoding has been inhibited to estimate savings thanks to storing intervals in the leaves and the binary topology. To inhibit the Huffman encoding, it has been forced an equiprobable statistic and the presence of 512 leaves in the tree so that it balances totally. So, population size and revocation % is not fixed. Obtained results are shown in Figure 3.

Notice that the difference increases when having realistic revocation % (around 10%). AD-MHT suffers when it has to answer many requests concerning valid certificates. That is due to the "adjacency checking". HuffMHT is insensible to this aspect (leaves store intervals).

Next we want to test HuffMHT scalability. To do so, population size varies while keeping revocation % fixed (realistic 10%). Statistic is equiprobable. Results can be seen in Figure 4.

Savings are greater when population size increases. It shows the scalability power of the HuffMHT implementation.

Next, we want to isolate the savings due to the Huffman codification. To do so, we define a parameter called Codification Gain (C_G) which refers to the savings (%) in BW_d obtained thanks exclusively to the encoding. To be more accurate, we also define a set of certificates of variable size whose serial numbers are more frequently asked. That frequency is also controlled and ranges from 20% to 50% of the requests. Obviously, this probability impacts on the effective % of requests concerning revoked certificates (i.e. when attacking the set in a 20%, effective % of requests concerning a revoked certificate is 28%).

Figure 5 shows that codification gain increases for any statistic different from equiprobable. On the other hand, a small set strongly frequently demanded leads to high C_G.

5 Conclusions

In this paper we have proposed an operative implementation of a system which manages to minimise the main constraint of an offline revocation system (bandwidth) maintaining a good performance in the rest of parameters (processing capacity) and all the advantages of an offline system (non-TTP distribution). System performance is what we have called balanced.

Huffman encoding allows to distribute leaves containing revocation information in a fashion which minimises the average response length. The rest of design decisions have been proved to be correct in order to reduce bandwidth. The HuffMHT provides high bandwidth savings for all kind of statistics which can be found in the status checking. The system performs better than balanced

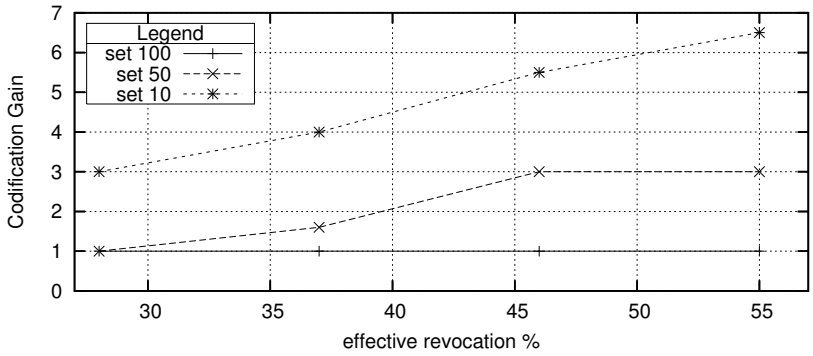


Figure 5: C_G obtained varying set size and revocation %

hash trees when the status of a subset of certificates is verified frequently. On the other hand, if status checking rates are similar for all certificates, our approach leads to a binary balanced tree which is the best option among the balanced trees. Besides, it has been proved that storing intervals in the leaves instead of single serial numbers leads to reduce bandwidth as well since "adjacency checking" is not necessary. HuffMHT scalability has also been proved to be high, what is essential for nowadays distributed environment.

Acknowledgements

This work has been supported by the Spanish Research Council under the project ARPA (TIC2003-08184-C02-02) and the European Research Council under the project UBISEC (IST-FP6 506926).

References

[1] R. Housley, W. Ford, W. Polk, and D. Solo. Internet X.509 Public Key Infrastructure Certificate and CRL Profile, 1999. RFC 2459.

[2] David Huffman. A method for the construction of minimum-redundancy codes. *IRE*, 40(9):1098–1101, 1952.

[3] ITU/ISO Recommendation. X.509 Information Technology Open Systems Interconnection - The Directory: Autentication Frameworks, 2000. Technical Corrigendum.

[4] P.C. Kocher. On certificate revocation and validation. In *International Conference on Financial Cryptography (FC98). Lecture Notes in Computer Science*, number 1465, pages 172–177, February 1998.

[5] R.C. Merkle. A certified digital signature. In *Advances in Cryptology (CRYPTO89). Lecture Notes in Computer Science*, number 435, pages 234–246. Springer-Verlag, 1989.

[6] J. Muñoz, J. Forné, O. Esparza, and M. Soriano. Certificate Revocation System Implementation Based on the Merkle Hash Tree. *International Journal of Information Security (IJIS)*, 2(2):110–124, 2004.

[7] J. Muñoz, J. Forné, O. Esparza, J. Pegueroles, and E. Pallares. Reducing the Communication Overhead of an Offline Revocation Dictionary. In *Trust and Privacy*, volume 3184 of *LNCS*, pages 269–278. Springer-Verlag, 2004.

[8] M. Myers, R. Ankney, A. Malpani, S. Galperin, and C. Adams. X.509 Internet Public Key Infrastructure Online Certificate Status Protocol - OCSP, 1999. RFC 2560.

[9] M. Naor and K. Nissim. Certificate Revocation and Certificate Update. *IEEE Journal on Selected Areas in Communications*, 18(4):561–560, 2000.

[10] ITU/ISO Recommendation X.509. Information technology Open Systems Interconnection - The Directory: Public Key and Attribute Certificate Frameworks, 1997.

Infraestructuras AAA en redes IPv6

Rafael Marín López, Ernesto Eulogio Blázquez , Gregorio Martínez Pérez, y Antonio F. Gómez Skarmeta
Departamento de Ingeniería de la Información y las Comunicaciones.
Facultad de Informática, Campus de Espinardo S/N
Universidad de Murcia,
30071 Murcia, España.
E-mail: {rafa, ernesto,gregorio,skarmeta}@dif.um.es

Abstract*. AAA (Authentication, Authorization and Accounting) frameworks are defined as a set of models, infrastructures and protocols needed to interconnect AAA entities. They have been mainly deployed so far using the RADIUS protocol in IPv4 networks. However, when taking such AAA systems to IPv6 networks some interesting issues appear that need to be addressed. This is the main focus of this paper where both RADIUS and Diameter protocols are deployed in different IPv6-related scenario.*

1. Introducción.

El mundo de las tecnologías está en continuo crecimiento. Sobre todo en una época en la que los dispositivos electrónicos son cada vez más avanzados, con más funcionalidad y con posibilidad de conexión sin cables habilitando a su usuario a poder moverse sin pérdida sustancial de conectividad permitiendo la movilidad del usuario entre distintas redes. PDAs, teléfonos móviles de última generación, PCs portátiles, incluso electrodomésticos que se conectan a Internet para comunicarse con su dueño que se encuentra en la oficina. Es difícil concebir equipos que de una u otra forma no necesiten conectividad.

Las organizaciones que dan servicios a todos esos dispositivos, conocidos como proveedores de servicios (SP), necesitan gestionar sus usuarios de una manera eficiente y segura. Cada vez es más común que los usuarios necesiten acceder a los servicios de su proveedor desde cualquier lugar del mundo, siendo muy habitual que un usuario que tiene una relación contractual con un SP intente acceder a servicios ofrecidos por otro SP, que tiene un acuerdo de negocio con el primero. Estas relaciones permiten lo que se de denomina *roaming* de usuarios entre diferentes dominios administrativos, concepto que cada vez cobra más importancia en la sociedad actual de las telecomunicaciones.

Para poder hacer frente a todas estas necesidades, se ha diseñado una infraestructura conocida como AAA que ofrece servicios de Autenticación, Autorización y de Accounting (Auditoría). Las infraestructuras AAA actuales están basadas en protocolos como *RADIUS* [11], los cuales fueron diseñados cuando las necesidades de estos servicios eran más simples. En la actualidad, estos protocolos han ido evolucionando de manera forzada para adaptarse con más o menos éxito a estas necesidades. Sin embargo, se estaba llegando a un punto en el que se hacía necesario un cambio drástico que solventara las carencias de esas infraestructuras y pudieran atender las nuevas necesidades. Con esos objetivos, nace un protocolo nuevo, *Diameter*, que pretende dar servicio AAA paliando las carencias de las infraestructuras actuales. Este nuevo protocolo se caracteriza por la extensibilidad y la adaptabilidad, siendo diseñado para poder realiza cualquier tipo de aplicación que se necesite y poder afrontar las necesidades de futuras tecnologías de acceso. Otra característica fundamental es que está totalmente orientado a la interconexión de diferentes dominios, característica básica en los entornos de trabajo actual.

Pero la meta de implantar el protocolo Diameter en las infraestructuras AAA no es un objetivo trivial. Los SP actuales tienen sus sistemas basados en otros protocolos como RADIUS o TACACS, por tanto no pueden desecharse tan a la ligera para cambiarlos por otros que soporten Diameter sin saber realmente si esto supone una gran mejora. Por otra parte, un cambio drástico de un protocolo a otro es una labor compleja. Por todo esto hay que ofrecer mecanismo y una guía para que los actuales SPs migren sus infraestructuras AAA a aquellas basadas en Diameter sin producir un gran impacto en el SP tanto a nivel económico como a nivel de infraestructura.

Con este fin, en este trabajo presentamos un análisis de las infraesrucuras AAA, en especial Diameter, y se realiza un estudio con detalle de escenarios de despliegue tanto en el caso de un dominio simple y el caso de múltiples dominios. Por último se especifican los diferentes servicios AAA que Diameter aporta a estas infraestructuras, centrándonos en las experiencias conseguidas dentro del proyecto europeo Euro6IX -European IPv6 Internet Exchanges Backbone-IST Project- el cual pretende fomentar la divulgación de IPv6 a nivel europeo mediante la creación de una gran red de servicios multidominio.

2. El Framework AAA

En este apartado se explica los conceptos involucrados en el acrónimo AAA, para pasar después a estudiar los protocolos y seleccionar uno de ellos como el más adecuado para ser utilizado en estas infraestructuras.

2.1. Autenticación, Autorización y Accounting

AAA define un framework donde poder aplicar procesos de autenticación, autorización y accounting. Intenta facilitar el proceso que permite estos servicios a través de las distintas redes y las distintas tecnologías que se pueden encontrar actualmente. Además, el uso combinado de estos tres procesos es un factor importante para lograr una gestión correcta y un control de los recursos y usuarios, y una seguridad adecuada en un dominio administrativo.

Para poder entender el concepto de AAA se debe entender bien cada concepto por separado.

Autenticación involucra la validación de la identidad de la entidad final para permitirle acceso a la red o a determinados servicios que un SP ofrezca. La entidad final presenta una credencial que lo identifica de manera unívoca. Ésta puede ser un login y password clásico, un certificado digital, una clave privada en una tarjeta inteligente, información biométrica (huellas digitales, información sobre retina ocular) y debe identificar al usuario.

Autorización se puede definir como los permisos que un usuario tiene para acceder a un determinado recurso o utilizar un determinado servicio. Autorización suele ir ligado a un proceso de autenticación previo pero hay que notar que se tienen que ver como procesos totalmente separados.

Auditoría o Accounting ofrece los medios necesarios para recoger información sobre el uso de los recursos que el usuario final ha utilizado. Esta información será del tipo, tiempo de uso, información enviada o recibida, etc. El SP utilizará esta información para temas referentes a estadísticas, estudios de utilización de recursos, pero sobre todo, para tema de tarificación y cobro.

2.2. Protocolos AAA

Dentro de los protocolos enmarcados en AAA se pueden destacar principalmente cuatro: TACACS[8], RADIUS[11], Diameter[1] y SNMP[10]. Actualmente la mayoría de entornos AAA están implementados utilizando RADIUS y en un menor porcentaje TACACS. Incluso puede haber entornos mixtos con varios tipos de protocolos actuando a la vez. Pero la situación actual ha cambiado mucho desde que se diseñaron estos protocolos, el volumen de usuarios de los entornos, las tecnologías de acceso, etc. Todo ha ido evolucionando y los requisitos que tiene que soportar las infraestructuras AAA son cada vez más.

El Internet Engineer Task Force (IETF), siguiendo un proceso de selección [12,13] basado en requerimientos de movilidad, NAS de nueva generación, y protocolos de roaming, determinó cual de los protocolos era el más adecuado para cumplir con las nuevas necesidades de servicios AAA. TACACS se desestimó inicialmente por las carencias que presentaba. RADIUS también presentaba serios problemas de extensibilidad, movilidad y seguridad entre otros. SNMP se consideró una buena alternativa para auditoria pero no estaba claro su papel en autorización y autenticación. Finalmente Diameter fue escogido como la principal alternativa pues ofrecía características para afrontar las nuevas necesidades AAA..

Así, se está potenciando el uso de este nuevo protocolo AAA no sólo en redes IP sino también en entornos de redes móviles para que sustituya las infraestructuras actuales y permita que estos entornos puedan satisfacer las necesidades de los SP.

El principal problema que se encuentra Diameter para conseguir una expansión rápida es que hay muy pocos dispositivos de red que lo implementen. La mayoría soporta RADIUS, por eso es tan importante la capacidad de convivencia que aporta Diameter mediante el uso de agentes. Los SP podrán ir migrando progresivamente a Diameter según sus dispositivos vayan soportando Diameter, permitiendo relacionarse unos dispositivos con otros.

3. Modelo de Infrastructuras AAA basada en Diameter.

En este apartado comenzaremos examinando las distintas entidades que Diameter aporta. Pasando después a estudiar las infraestructuras AAA Diameter de dominio simple, donde se tratan temas referentes a la organización interna dentro del propio dominio de la organización, y de múltiples dominios, donde cobran importancia los aspectos de la infraestructura relacionados con la interacción con otros dominios. Por último, se revisan los servicios AAA que Diameter ofrece a estas infraestructuras.

3.1. Entidades Diameter

En su especificación base [1], el protocolo Diameter define varias entidades AAA que podrían ser introducidas en una infraestructura AAA. Estas entidades tienen diferentes funciones como enrutar información AAA (agentes *Relay* o agentes *Proxy*); proveer de información de enrutamiento AAA para alcanzar otras entidades AAA (agente *Redirect*); traducir paquetes RADIUS a paquetes Diameter y viceversa (agente *Translator* o de traducción) y finalmente servidores AAA que almacenan la información relacionada con un cliente para el proceso de verificación de credenciales. En general, a todos los agentes se les conoce con el nombre de brokers Diameter.

Por tanto se dispone de estos elementos para construir infraestructuras AAA generales. Para hacerlo, dos casos deben tenerse en cuenta: el desarrollo de una infraestructura AAA considerando un único dominio (caso *single domain*) y la interconexión de estas infraestructuras AAA (que pertenecen a diferentes dominios administrativos) para proveer roaming de usuarios entre dominios (caso *multi-domain.*)

3.2. Caso Dominio Simple

Las infraestructuras AAA de un único domino tienen una combinación de entidades Diameter y un conjunto de relaciones entre ellas para realizar transacciones AAA. En la Fig.1.a se muestra una interacción típica de esas entidades dentro de una infraestructura de un único dominio.

En este modelo tenemos numerosas alternativas (fig.1.a). NASs pueden directamente enviar peticiones AAA a un servidor Diameter o enviar las peticiones a través de elementos intermedios como agentes relays o agentes proxy. La última opción sucede normalmente cuando los NASs están geográficamente muy lejos del servidor Diameter o cuando hay muchos NASs distribuidos en la red y no se quiere tener muchas conexiones directas entre NAS y servidor Diameter, evitando así problemas de escalabilidad. En ocasiones los agentes y los servidores Diameter podrían necesitar información de enrutamiento AAA para transportar peticiones AAA a determinados servidores Diameter. También es posible que los NAS utilicen clientes AAA basados en protocolos antiguos como RADIUS. En este caso es necesario desplegar agentes de traducción para dar soporte a este tipo de NAS. Además, toda la información intercambiada entre los distintos agentes y servidores Diameter debe ser securizada de alguna manera ya que pueden contener información sensible como claves, secretos compartidos, información personal, información de accounting, etc. Para solucionar este problema, Diameter define una seguridad salto a salto basada en IPSec (obligatoria) o en TLS (opcional).

3.3. Caso dominio múltiple

En ocasiones, servidores Diameter (SD) reciben información AAA de usuarios finales que pertenecen a un dominio diferente (usuarios en roaming). En este caso, los SD no tienen suficiente información para ejecutar procesos sobre el usuario final. Para solucionar esto, los SD tienen que reenviar la información AAA a la infraestructura AAA del dominio local (*home domain*) del usuario final para procesarla y tomar una decisión sobre los procesos de autenticación y autorización. Esto es un claro ejemplo de como una infraestructura local AAA necesita interactuar con otras infraestructuras AAA situadas en otros dominios para permitir roaming de usuarios finales.

Como se puede ver en la Fig.1(b,c), se pueden establecer dos puntos de vista distintos sobre las relaciones entre los diferentes dominios administrativos. Por un lado la Fig.1.b muestra un modelo semi-jerárquico donde se definen relaciones de dependencia entre distintos dominios. Los dominios en la parte superior de la jerarquía ofrecen servicios a los dominios de los niveles inferiores. Por ejemplo, dominios nacionales que dan servicio a dominios regionales. En este esquema, también son posibles relaciones de dependencia entre dominios del mismo nivel (Por ejemplo, un dominio nacional que da servicio a otro domino nacional).

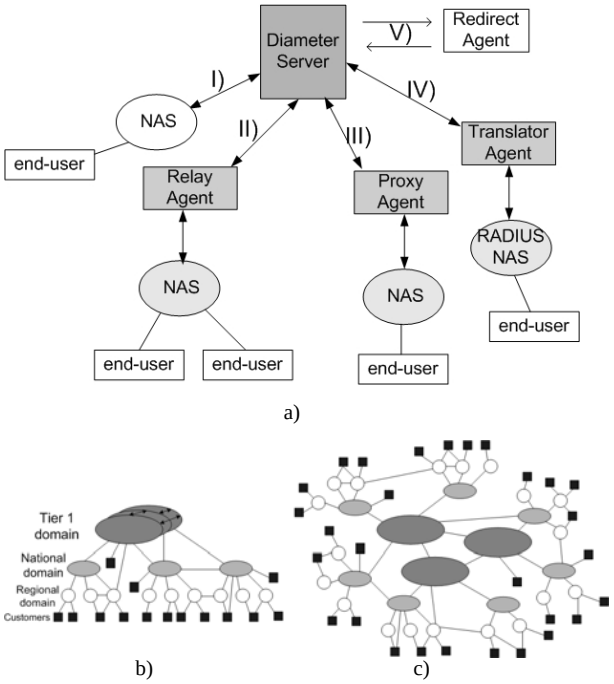


Fig. 1. Caso dominio simple (a) –Vistas de dominio múltiple (b,c).

Por otro lado, la Fig.1.c muestra relaciones sin dependencia entre dominios. Estas relaciones son necesarias para permitir a los usuarios finales que pertenecen a un dominio administrativo moverse a un dominio foráneo (provisión del servicio de roaming). En ambas casos, las relaciones están basadas en acuerdos de negocios que se materializan mediante la interconexión de las infraestructuras AAA desplegadas en cada dominio administrativo.

Para la gestión de roaming es necesario transportar información de autenticación y autorización de un dominio foráneo (donde el usuario final realiza el roaming) al dominio local del usuario como se describe en [2]. Esta información puede ser enviada a través de una infraestructura basada en diferentes tipos de interconexiones de agentes Diameter o brokers (Agentes relay, proxy, redirect o de traducción). Todos podrían formar parte de un *Backbone Broker AAA Diameter*¸es decir un conjunto de agentes interconectados que a su vez permiten la conexión de dos dominios (ver Fig. 2).

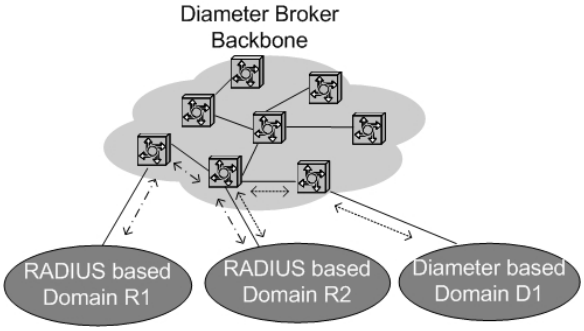


Figura 2: Backbone Broker AAA Diameter

Destacar que en este tipo de medios con elementos intermedios, RADIUS tiene numerosos problemas como ya se ha comentado, debido a que la especificación del protocolo no define

comportamiento específico de agentes ni requiere una seguridad salto a salto fuerte. No obstante, TERENA[14] propone una solución basada en el uso de DNSsec[16] para obtener las credenciales del nodo con el que se quiere conectar. Estas credenciales se utilizarían para establecer el secreto compartidos con el que crear una conexión segura. Esta conexión sería directa, sin brokers intermedios. Diameter, por su parte, obliga en su especificación a establecer túneles IPSec (o TLS) entre las entidades Diameter que se conectan directamente.

Decir que para Diameter se diseñó una aplicación con un módulo llamado CMS [19] que ofrecía seguridad extremo a extremo, evitando así que los elementos intermedios o brokers pudieran ver y modificar la información contenida en los mensajes. Pero se desestimó al encontrarse deficiencias en el diseño que impedían cumplir con el objetivo que tenía dicho módulo.

Una solución (que aún está en desarrollo) es usar agentes de redirección que podrían informar sobre la manera de alcanzar un determinado servidor Diameter en otro dominio. Dicho agente enviaría la información de enrutamiento a la entidad que la solicita añadiéndole una credencial de seguridad de la entidad AAA remota con la que quiere comunicarse. De este modo, Diameter permitiría el flujo de información de manera segura entre distintos dominios administrativos sin pasar por elementos intermedios. Uno de los principales problemas de esta solución es que requiere el uso de una infraestructura de PKI en cada dominio. Además, es necesario definir relaciones de confianza entre las PKI's para la validación cruzada de certificados. Una propuesta interesante es utilizar un servidor DNSsec en lugar de una PKI para obtener las credenciales, tal como se hace en Opportunistic Encryption[15].

4. Proyecto Euro6IX

El proyecto Euro6IX -European IPv6 Internet Exchanges Backbone-IST Project [3] – es una iniciativa europea con dos objetivos principales. El primero es el de proporcionar la base para la introducción rápida de IPv6 en Europa mediante estudios avanzados en diseño y desarrollo de redes, así como el desarrollo y portabilidad de servicios y aplicaciones avanzadas finalmente validadas por grupos de usuarios en pruebas internacionales. El segundo objetivo está relacionado con la divulgación activa de resultados y experiencias obtenidas de la actividad de esta investigación.

La aproximación técnica seguida para obtener ambos objetivos está basada en la construcción de un gran backbone escalable y nativo IPv6 de intercambio de tráfico (IX) proporcionando conectividad a través de Europa. Por encima de todo, y para promover el uso de IPv6, servicios claves y aplicaciones están siendo desarrolladas o migradas al nuevo protocolo. Todo esto permitirá conseguir experiencias IPv6 reales envolviendo en el proceso a usuarios finales, y compañías Telco e ISPs.

En concreto, en el proyecto Euro6ix uno de los principales objetivos es soportar roaming así como control de acceso a la red. La principal razón es el aspecto multi-domino de la red Euro6ix. Es decir, cada IX define un gran domino administrativo ofreciendo servicios a clientes. Éstos pueden ser usuarios finales o incluso otros dominios administrativos dependientes de él.

Tradicionalmente, un IX solo ofrece servicios de red de nivel 2. Sin embargo la red Euro6IX concibe un IX con la capacidad de proveer servicios de red más altos y avanzados (nivel 3, nivel 4,…). Para controlar y gestionar estos servicios ha aparecido la necesidad del despliegue de infraestructuras AAA para realizar de forma segura y eficiente estos servicios. Este escenario tiene algunas características más de interés:

- Algunos IXs están directamente interconectados pero otros se conectan mediante varios IX intermedios. En otras palabras, la información entre dos dominios administrativos podría pasar por otro dominio intermedio. Bajo este escenario, se necesita seguridad extremo a extremo para algunas transacciones.

- Después del análisis previo la infraestructura AAA de Euro6IX debe estar basada en Diameter como protocolo principal.

- Para ofrecer material criptográfico, se deben desplegar infraestructuras de clave pública (PKI) en cada IX para certificar a las diferentes entidades en cada domino.

- Se establecen asociaciones entre PKIs pertenecientes a diferentes dominios (IX) mediante diferentes mecanismos (certificación cruzada, *bridge-CA*). Gracias a ello, un certificado entregado por una PKI en un dominio será de confianza para otra PKI en otro dominio distinto y viceversa.

Finalmente, hay que destacar que la red Euro6IX hace frente a usuarios finales heterogéneos (usuarios inalámbricos y fijos) y tecnologías de acceso como xDSL, wireless LAN, etc. Bajo esas condiciones se proponen numerosas recomendaciones de diseño para infraestructuras AAA en Euro6IX. Esas recomendaciones han sido divididas en dos partes: desarrollo interno de infraestructuras AAA en cada IX (Dominio simple Euro6IX) e interconexión de esas infraestructuras para soportar roaming en la red (Dominios Múltiple Euro6IX).

4.1. Dominio simple Euro6IX

Cada organización tiene requisitos y políticas de gestión que imposibilitan imponer una infraestructura concreta. Por eso se han elaborado diversas

recomendaciones para desarrollar infraestructuras AAA dentro del proyecto Euro6IX

La primera recomendación es que la infraestructura debe estar basada en el protocolo Diameter. Gracias a ello se pueden considerar diferentes agentes Diameter y pensar cuales pueden ser desarrollados. Para decidir que agentes utilizar debemos considerar algunos aspectos que exponemos a continuación.

Un IX gestiona usuarios finales. Por esto debe considerarse como un "macro" proveedor de servicios que tiene que gestionar información AAA de usuarios finales. Será necesario al menos un servidor Diameter con información de usuarios.

La información AAA generada por usuarios finales debe ser enviada entre servidores de acceso a la red (NAS) y servidores Diameter. Por cuestiones de escalabilidad tal y como mencionábamos en el apartado 3.2 y puesto que el domino definido por un IX puede ser grande y numerosos NAS pueden dar servicio a muchos usuarios finales es necesario la introducción de servidores intermedios que reciban información AAA de un conjunto de NASs y reenviar la información hacia los servidores Diameter. En este sentido, el conjunto de asociaciones de seguridad gestionados por el servidor Diameter es reducido (solo las que tenga con los servidores intermedios).

En esta aproximación, estamos considerando que la seguridad extremo a extremo entre NASs y servidores Diameter no es necesaria porque los elementos intermedios pueden considerarse de confianza. Es un hecho razonable teniendo en cuenta que dentro de un IX todos los dispositivos están desarrollados y gestionados por la misma organización y, además, es posible establecer seguridad salto a salto donde se considere oportuno.

Con respecto a los elementos intermedios mencionados antes, hay dos alternativas: agentes relay o agentes proxy como vimos en el apartado 3.2. Se recomienda usar la primera opción, debido a que no son necesarias las aplicaciones de políticas que los servidores proxy ofrecen. Los agentes relay enrutan de manera más rápida ya que no hacen ningún procesamiento del mensaje. En cualquier caso, si un dominio debe aplicar políticas tiene que reemplazar los agentes relay por agentes proxy.

En resumen, en un dominio AAA simple en Euro6IX la infraestructura estaría constituida con la opción de II de la Fig.1.

4.2. Dominio múltiple Euro6IX

Cada IX y sus recursos están gestionados por diferentes organizaciones, por lo que son considerados como diferentes dominios administrativos. Incluso ofrecen servicios y recursos a otros SPs aparte de a sus usuarios finales.

Por otro lado, cada SP define otro domino de red administrativo con sus propios recursos, servicios y usuarios finales. Podría utilizar también recursos de un IX. Esto es, coincide con la visión jerárquica mostrada en la Fig.1.b). La red de la Fig.1.c) también es posible en el sentido de que ambos dominios (SP y IX) podrían no establecer relaciones de dependencia entre ellos. Ambas vistas muestran IX con diferentes roles.

En el primer caso IX es un dominio administrativo con sus propios usuarios finales (rol de SP a usuarios finales) que pueden establecer acuerdos con otros dominios administrativos como aquellos definidos por SPs. Estos acuerdos definen relaciones igual a igual y se realiza conectando la infraestructura AAA del dominio local del IX con la infraestructura AAA del SP. Ambos dominios SP e IX están considerados en el mismo nivel desde un punto de vista lógico. Usuarios finales pueden hacer roaming al dominio IX y viceversa.

En el segundo caso, los SPs podrían necesitar recursos ofrecidos por un IX (rol de SP), y entonces se establecería una relación de dependencia entre IX y SP. Así se puede decir que esos SPs pertenecen al macro-dominio definido por IX y que están bajo su control.

De este modo, si un IX está realizando un rol de SP, permitirá que la información AAA generada por un dominio dependiente llegue a otro domino remoto para dar soporte de roaming de usuarios finales entre ambos dominios. Esta interconexión se puede realizar de dos modos: estableciendo una conexión directa entre los servidores AAA de ambos dominios o estableciendo una comunicación indirecta mediante un backbone de brokers. En la red Euro6IX la opción elegida es la del Backbone AAA. Éste está construido conectando las infraestructuras AAA de los IXs para dar servicios AAA a otros dominios (infraestructura AAA Brokering) con otras infraestructuras AAA Brokering desarrolladas por otros dominios de IX. De manera esquemática podemos verlo en la Fig.3.

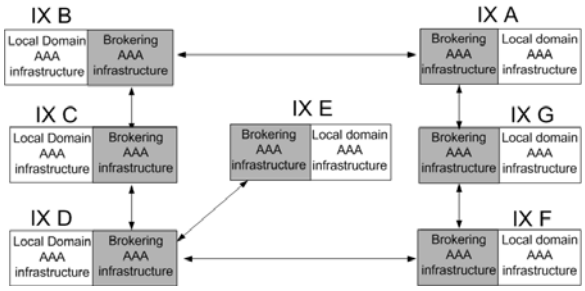


Figura 3: Dominio múltiple en Euro6IX

Como se pude ver, cada infraestructura AAA de IX está conectada con un par de infraestructuras pertenecientes a otros dominios. Además hay establecidas asociaciones de seguridad (SA) con las infraestructuras AAA de los IX adyacentes. Esta asociación de seguridad se implementa usando el material criptográfico ofrecido por la PKI de cada IX.

4.3. Servicios AAA en la red Euro6IX

Como se ha mencionado, la infraestructura AAA Brokering da un conjunto de servicios AAA que dependen de los dominios SPs. En el análisis inicial se veía que se podían ofrecer varios servicios (enrutamiento AAA, información de routing AAA y traducción AAA). En Euro6IX se dan los tres servicios de la siguiente manera.

- Servicio de Enrutamiento AAA Euro6IX: se necesita transportar información de un SP a otro SP bajo control del IX. Entre Agentes Relay y Proxys la elección es implementar agentes Proxys. La razón principal es que los acuerdos entre dominios hay que llevarlos a cabo y estos agentes permiten realizar dicha tarea.

- Servicio de Información de Enrutamiento AAA Euro6IX. Realizado por los agentes de redirección. Los IX gestionan la información necesaria para alcanzar un determinado dominio, incluso si este domino está gestionado por otro IX. Esto se debe principalmente a que se tiene una visión global de la red ya que se puede saber qué dominios están bajo su control y además se puede obtener información de los dominios que están bajo el control de otro IX.

- Servicio de Traducción AAA Euro6IX. Un requisito de Euro6IX es ofrecer soporte a sistemas de dominios basados en protocolos como RADIUS. Esto permite a las redes incluir dominios basados en protocolos RADIUS y permitir una transición progresiva a sistemas Diameter. Hay que notar que algunos participantes del proyecto comenzaron a desarrollar escenarios basados en RADIUS debido a su falta de familiaridad con Diameter y a los que hay que dar la posibilida de integración en la nueva infraestructura AAA basada en Diameter.

En resumen, la infraestructura AAA brokering debe implementarse con agentes proxy para ofrecer servicios de enrutamiento, agentes redirect para ofrecer información de enrutamiento y agentes de traducción para permitir servicios de traducción RADIUS/Diameter.

5. Escenarios de Despliegue

A continuación se muestra un conjunto de escenarios de tests que muestran la evolución desde RADIUS a Diameter en el proyecto de Euro6IX.

5.1. Escenario RADIUS-Diameter

Debido a las limitaciones que el protocolo RADIUS presenta para la gestión de escenarios de roaming y escenarios multidominio, el protocolo Diameter ofrece una mejor opción para dar soporte a diferentes tipos de accesos de red a Euro6IX, como es el caso de la gestión de usuarios y dispositivos móviles. Sin embargo, el protocolo RADIUS es ampliamente utilizado por TELCOs e incluso por los desarrolladores que están creando las pruebas de IPv6 y AAA.

De este modo, se prefiere una transición suave de las infraestructuras basadas en RADIUS a infraestructuras basadas en Diameter.

Esta transición consiste principalmente en el desarrollo del núcleo de una infraestructura AAA basada en Diameter y permitir a otros dominios acceder a esa infraestructura núcleo sin que tengan que cambiar su propia infraestructura AAA (que puede estar basada en RADIUS, por ejemplo). Para conseguir este objetivo, NASREQ[4] define la funcionalidad de un traductor RADIUS-Diameter. Siguiendo [4], se ha desarrollado un traductor básico.

La Fig.4 muestra el escenario utilizado para probar el traductor. Como se pude ver, está basado en la delegación de prefijos DHCPv6. En este escenario el cliente DHCPv6 (un router 6WIND 6100) envía una petición al servidor DHCPv6 (router 6WIND 6200) para conseguir un prefijo IPv6 que será utilizado por el cliente DHCPv6 para enviar Router Advertiments con ese prefijo a través de otra interfaz donde los usuarios finales está conectado. El servidor DHCPv6 examina el identificador (DUID) del cliente DHCPv6 y envía una petición RADIUS al servidor AAA con un login y un password que el propio servidor DHCPv6 almacena para ese cliente. El servidor RADIUS verifica ambos elementos y si la autenticación es correcta, el prefijo será devuelto al servidor DHCPv6. El servidor que se ha instalado es un servidor Diameter implementando parte de [4] que contiene la información del prefijo asociado a un determinado login y password. Para que el cliente RADIUS puede interactuar con el servidor Diameter es necesario la introducción de un traductor RADIUS/Diameter. Éste capturará la solicitud RADIUS, creará una petición Diameter y se le enviará al servidor Diameter NASREQ. Cuando el servidor responda, el traductor recibirá dicha respuesta y la convertirá a RADIUS, enviándosela al servidor DHCPv6 de vuelta. Hay que destacar la transparencia total de este proceso de traducción para el cliente RADIUS que de hecho, no sabe que está hablando con un servidor Diameter.

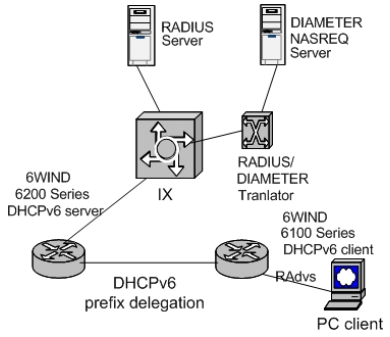


Figura 4: RADIUS-Diameter transition scenario

Como nota final decir que el traductor se ha implementado utilizando parte del código de un servidor RADIUS de libre distribución freeRADIUS [18] y para la parte de Diameter utilizando la implementación *Open Diameter* [7]

5.2. Escenario Diameter

Este escenario (ver Fig.5), está basado únicamente en el protocolo Diameter. Se utiliza IEEE 802.1X como tecnología de acceso a la red, mientras que se usa EAP-TLS [5] para la autenticación. Los usuarios pueden utilizar sus certificados X.509 para acceder a la red.

Para crear este escenario se ha modificado la implementación HostAP[6] añadiendo un cliente Diameter. Hay que indicar que la implementación Diameter utilizada en este escenario también es Open Diameter el cual soporta IPv6. Además, se han implementado los métodos EAP-TLS integrados en el servidor Diameter. También, se han desarrollado los métodos EAP-TLS en la parte del servidor y se han integrado en el servidor Diameter usando también la implementación de Open Diameter.

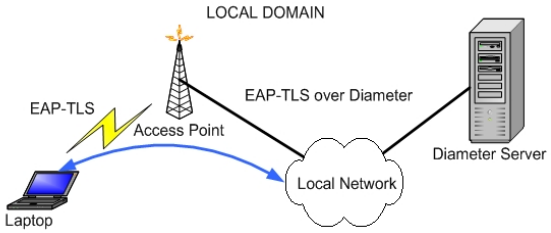


Figura 5: Escenario Diameter con 802.1X y EAP-TLS

De este modo, estos cambios permiten autenticar usuarios finales sobre Diameter en lugar de RADIUS, pero los usuarios no percibirán ningún cambio en el proceso de autenticación, requisito principal de este escenario.

5.3. Escenario Diameter Multidominio

El escenario completo consta de tres dominios. A, B y C. Todas las máquinas de todos los dominios están montadas sobre Open Diameter. Es decir, todos los servidores o bien son servidores Diameter o bien están utilizando una aplicación que corre sobre Diameter. Todo el escenario está montado sobre IPv6.

Una visión general del escenario (Fig.6) sería la siguiente. Un usuario (un nodo móvil) que se desplaza a un dominio que no es el suyo y quiere tener acceso a él. Se pretende autenticar usando EAP-TLS. Se conecta a la red del dominio foráneo usando PANA[9] para transportar la información EAP hasta el NAS. Un vez que el NAS tiene dicha información, este debe conectar con el servidor EAP del dominio local del usuario. Una vez que el cliente consiga autenticarse podrá tener acceso a los servicios del dominio a los que esté autorizado. Para que el agente de redirección entre en uso, el domino A no conoce el dominio local (C) del cliente. Así, debe obtener de alguna manera la información necesaria para contactar con dicho domino. Esta información se la ofrecerá el agente de redirección. Como ya se ha comentado, cuando el servidor no sepa como enrutar información hacia un dominio, este le preguntará al agente de redirección, quien responderá indicándole al servidor por donde debe encaminar dicha información para alcanzar el dominio deseado. La información de enrutamiento que tiene el agente de redirección refleja los acuerdos que hay entre el dominio local y el resto de dominios. Si entre dos dominios no hay acuerdo, el agente de redirección no tendrá información de enrutamiento hacia dicho dominio y, por tanto, el servidor no podrá enrutar a ese dominio.

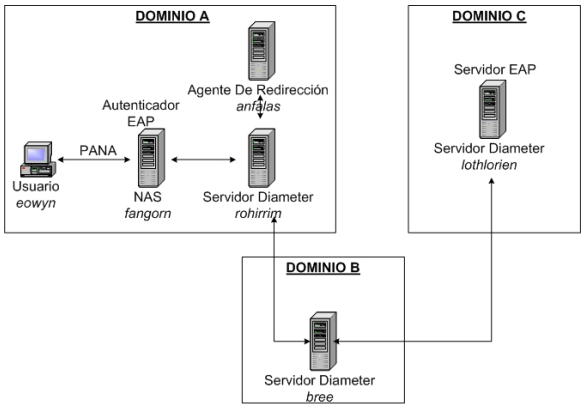


Figura 6: Escenario Diameter multidominio

6. Conclusiones

En este trabajo se han presentando las principales características y escenarios de despliegue de las infraestructuras AAA basadas en el protocolo Diameter en especial en el contexto de las redes IPv6, a través de los casos evaluados en el proyecto europeo Euro6IX. Se han detallado tres escenarios concretos de diverso tipo. El primero ha sido un escenario heterogéneo donde conviven simultáneamente el antiguo protocolo RADIUS con el protocolo Diameter. En concreto está basado en la delegación de prefijos IPv6 y en el se ha mostrado el uso de un traductor entre RADIUS y Diameter. El segundo escenario es puramente Diameter. En él se ha utilizado IEEE 802.1X para acceder a la red y EAP-TLS para autenticar. El último escenario está también basado en Diameter, pero al contrario que en los otros dos escenarios, encontramos interacción entre varios dominios administrativos distintos.

Agradecimientos

Este trabajo está parcialmente soportado por los proyectos Euro6IX project (IST 2001-32161) y SAM TIC2002-04531-C04-04

Referencias

[1] Calhoun P., Loughney J. “Diameter base protocol”, RFC 3588, September 2003

[2] Charles E. Perkins, “Mobile IP Joins Forces with AAA” IEEE Personal communications August 2000

[3] Euro6IX EU IST Project, "European IPv6 Internet Exchanges Backbone", http://www.euro6ix.org.

[4] CalhounP., ”Diameter Network Access Server Application”, Draft July 2004.

[5] B.Aboba, J.Simon, “PPP EAP TLS Authentication protocol”, RFC 2716, October 1999

[6] Host AP driver for Intersil Prism2/2.5/3 and WPA Supplicant, http://hostap.epitest.fi/

[7] Open Diameter. http://www.opendiameter.org/

[8] Solar Designer “An analysis of the TACACS+ protocol and its Implementations” May 2000.

[9] Y. Ohba, A. Yegin, “Protocolo for Carrying Authentication for Network Access(PANA)” IETF DRAFT, October 2004.

[10] W.Stalling, “SNMPv3: A Security Enhancement for SNMP”, IEEE Comunication Surveys. Four Quarter 1998

[11] Rigney C., Willens S. “Remote Authentication Dial In User Service (RADIUS)”, RFC 2865, June 2000.

[12] Aboba B., Calhoun P. “Criteria for Evaluating AAA Protocols for Network Access”, RFC 2989 Nov. 2000

[13] Mitton D. ”Authentication, Authorization, and Accounting: Protocol Evaluation” RFC 3127, June 2001

[14] H. Eertink, A. Peddemors, R. Arends, K. Wierenga‡ “Combining RADIUS with Secure DNS for Dynamic Trust Establishment between Domains” TERENA, 2005.

[15] M. Richardson, D. Redelmeier, “Opportunistic Encryption using The Internet Key Exchange”, Internet-Draft, January 2005.

[16] R. Arends, R. Austein, M. Larson “DNS Security Introduction and Requirements”, RFC 4033, March 2005.

[17] Pat R. Calhoun, Stephen Farrell,”Diameter CMS Security Application”, Internet-Draft, March 2002.

[18] FreeRADIUS server, http://www.freeradius.org

Estudio experimental de los protocolos IP en redes inalámbricas multi-salto basadas en el protocolo DSR

Ramón Agüero, Johnny Choque, Jorge Lanza, Luis Sánchez y Luis Muñoz
Grupo de Ingeniería Telemática. Universidad de Cantabria
ETSII y de Telecomunicación. Avda los Castros s/n.
39005 - Santander (Cantabria)
Teléfono: 942 20 13 92 (Ext 14) Fax: 942 20 14 88
E-mail: [ramon, jchoque, jlanza, lsanchez, luis]@tlmat.unican.es

Abstract *The importance of multi-hop communications in the forthcoming wireless personal communication scenarios is continuously increasing. Although they have gathered a significant interest from the research community in the latest years, there are still a large number of challenges that need to be tackled. One of the most important aspects is that there is not much real experimentation with such protocols. Most of the studies are based on simulation analysis and therefore it is really difficult to see how scalable and feasible are the different solutions from an experimental point of view, especially considering the characteristics of existing technologies, which were not originally designed to cope with multi-hop topologies, such as IEEE 802.11 family of standards. The final paper will target this shortage, providing a fully experimental characterization of IP-based communications over real multi-hop platforms, deployed by means of a complete Dynamic Source Routing protocol implementation developed at the University of Cantabria, which includes, amongst others, the functionality of connectiong an ad hoc network to the external world, by means of a* ***gateway*** *entity.*

1. Introducción

A pesar de que en la actualidad la mayoría de las comunicaciones inalámbricas se basan en el *paradigma del último salto*: entre un teléfono móvil y una estación base en las comunicaciones celulares o entre un terminal y un punto de acceso en aquellas basadas en redes de área local inalámbricas o WLANs (*wireless local area networks*), se prevé que la presencia de las comunicaciones multi-salto sea mayor en un futuro no muy lejano.

En contraposición a estos escenarios más tradicionales, las redes *ad hoc* surgen como alternativa para poder habilitar las comunicaciones en aquellas situaciones en las que no se dispusiera de infraestructura subyacente. En este sentido, los ejemplos que más se han venido empleando son situaciones de emergencia (tras desastres naturales) o bélicos. Sin embargo, es posible que la relevancia comercial de ellos sea escasa, pero estas topologías estarán demandadas por los futuros escenarios de las comunicaciones móviles, que están centrados en el usuario como elemento principal, en lo que se ha venido a denominar como *Systems Beyond 3G* (B3G) [1]. En este sentido, tanto los operadores tradicionales como los nuevos creen que este tipo de topologías (las multi-salto) pueden emplearse para aumentar la cobertura en ciertas zonas (incluso de manera temporal) o para incrementar la capacidad de los accesos ya existentes, de manera eficiente en lo que se refiere al coste asociado, tal y como se desprende en [2].

Sin embargo, la existencia de validaciones reales de protocolos que habiliten este tipo de topologías es escasa. Aunque se lleva cierto tiempo trabajando en las labores de estandarización, coordinadas desde el grupo de trabajo MANET del IETF, se requiere que las diferentes soluciones, además de que se validen mediante simulación, se prueben en entornos reales, con el fin de comprobar su escalabilidad y validez desde un punto de vista eminentemente experimental.

2. Descripción del protocolo DSR

2.1. Introducción

El protocolo DSR (*Dynamic Source Routing*) pertenece a la familia de los protocolos de enrutamiento para redes *ad hoc* reactivos, en tanto y cuanto no requiere de la propagación periódica de mensajes de señalización con información topológica de la red. La característica que diferencia DSR frente a otras alternativas reactivas, como AODV (*Ad Hoc On Demand Distance Vector*), es que utiliza un esquema de enrutamiento fuente, y todos los datagramas llevan información completa acerca de la ruta que tienen que seguir hasta llegar a su destino. Tiene dos mecanismos principales: el descubrimiento de ruta y el mantenimiento de ruta. Además, su especificación [3] ha ido evolucionando y ha incorporado un gran número de aspectos adicionales.

2.2. Descubrimiento de Ruta

El descubrimiento de ruta es el mecanismo que se emplea para encontrar un camino hacia un destino. Cuan-

do una estación genera un datagrama, busca una ruta válida hacia el destinatario; si no la tuviera, difundiría (*broadcast*) un paquete *Route Request* (RREQ) que será recibido por todos los nodos que están dentro de su área de cobertura (vecinos). El RREQ se va propagando por la red, hasta que llega al nodo destino -objetivo- que responderá con un paquete *Route Reply* (RREP).
De esta manera, al recibir un RREQ, se pueden dar tres posibles situaciones:

- El nodo es el destino del RREQ y, por tanto, envía al origen un paquete RREP, en el que le indica, al origen, la ruta completa.
- Ya se había recibido un RREQ idéntico (mismo origen, mismo destino y mismo identificador), por lo que se descarta el paquete; esta situación se produce por la naturaleza *broadcast* del medio inalámbrico.
- Si no se cumple ninguna de las condiciones anteriores, el nodo añade su propia dirección a la ruta que viaja en el RREQ y vuelve a difundir el mismo.

Como se puede ver, cuando un RREQ llega al destino, este lleva consigo la ruta completa que ha atravesado hasta alcanzarle. Para enviar el RREP al origen, este debería buscar una ruta válida y, en caso de no encontrarla, iniciar un nuevo proceso de descubrimiento. En algunas ocasiones, cuando los enlaces inalámbricos se pueden considerar bidireccionales (es decir, cuando la comunicación en un sentido implica la del opuesto), el nodo destino puede invertir la ruta que recibe en el RREQ para transmitir el RREP.
Desde el momento en el que inicia el proceso de descubrimiento de ruta, el nodo origen debe guardar todos los datagramas dirigidos al nodo destino en un *buffer* o memoria local y en el momento de recibir el RREP procederá a su transmisión, utilizando la ruta reportada. En caso de no recibir el RREP en un intervalo definido, el nodo procederá a la retransmisión del RREQ, hasta un número máximo de veces, tras el cual descartará los paquetes almacenados en el *buffer*. El tiempo que transcurre entre transmisiones consecutivas de RREQ con el mismo nodo destino se va doblando paulatinamente, según un algoritmo *backoff exponencial binario*.

2.3. Mantenimiento de Ruta

Una vez que el nodo fuente es consciente de una ruta válida hacia el destino, empieza con la transmisión de la información. Para ello sitúa en cada uno de los datagramas, la ruta completa que tiene que seguir, de manera que los nodos intermedios que procesan el paquete, simplemente tienen que ver cuál es el siguiente nodo, para enviárselo. El proceso de mantenimiento de ruta se utiliza para comprobar que ésta sigue siendo válida mientras se siga usando. Para ello, cada nodo comprueba que el paquete llegue correctamente al siguiente salto de la ruta. En caso de que un enlace se rompa y un nodo lo detecte, mandará un paquete de *Route Error* (RERR) a la fuente, para que inicie un nuevo proceso de descubrimiento.
Se especifican tres posibles mecanismos para que los nodos aseguren la llegada correcta de un datagrama al siguiente salto:

- **Reconocimiento Pasivo.** En este caso, cada nodo permanece a la escucha tras la retransmisión de un paquete. Si este es retransmitido de nuevo por el siguiente salto, llega a la conclusión de que llegó correctamente al mismo. Este mecanismo tiene dos problemas principales: por un lado, no se puede emplear en el caso de que el siguiente salto sea el destino final del datagrama; además, requiere que los dispositivos inalámbricos trabajen en modo *promiscuo*, para poder escuchar los datagramas transmitidos por otros nodos, lo que supone un mayor gasto energético.
- **Reconocimientos de Nivel de Enlace.** Si la tecnología subyacente proporcionara una entrega con confirmación, el procedimiento de mantenimiento de ruta debería aprovecharla para detectar la posible caída de enlaces. De esta manera se limita la sobrecarga adicional impuesta por otras alternativas.
- **Reconocimientos DSR.** Si ninguna de las alternativas anteriores fuera factible, el protocolo DSR especifica la utilización de mecanismos de reconocimientos propios de DSR, mediante dos campos específicos en su cabecera. Esta tercera alternativa, a pesar de ser la más conservadora, lleva implícita cierta sobrecarga adicional.

2.4. Mecanismos Adicionales

Adicionalmente al comportamiento básico que se ha descrito anteriormente, la especificación de DSR recoge un conjunto de mejoras opcionales, que permiten mejorar el comportamiento del protocolo original. Estas se definen tanto en el proceso de descubrimiento como en el de mantenimiento de ruta; como no se han incorporado (por diferentes razones) a la implementación que se caracterizará posteriormente, no se describirán en detalle, sino que simplemente se enumeran a continuación:

- **Descubrimiento de Ruta.** Hay tres posibles mejoras adicionales a lo descrito con anterioridad: *Utilización de información topológica escuchada*, *Respuesta a RREQ con información local*, y *Limitación de los saltos en el RREQ*.
- **Mantenimiento de Ruta.** En este caso se definen hasta cuatro mecanismos adicionales que permiten mejorar el comportamiento original del proceso de mantenimiento de ruta: *Salvado de paquetes*, *Encolado de paquetes dirigidos a través de un enlace roto*, *Acortamiento automático de la ruta* y *Aumento de la difusión de los RERR*.

- **Extensión opcional Flow State.** Uno de los principales problemas que tiene la especificación original de DSR es la alta sobrecarga que supone la inclusión, en cada uno de los paquetes de datos de toda la ruta que tiene que seguir hasta alcanzar el destino. El problema se agrava más si cabe al emplear la nueva versión del protocolo IP (IPv6), en el que las direcciones son sensiblemente más largas. Con el objetivo principal de resolver este problema se incorporó la extensión *Flow State*, en la que en lugar de incorporar la ruta completa en cada paquete, se manda un identificador unívoco para ella.

3. Implementación del protocolo

3.1. Marco de desarrollo: Netfilter

El marco de desarrollo *Netfilter* [4] se trata de una facilidad ofrecida a partir de las versiones 2.4.x del *kernel* de Linux, que permite el procesamiento de paquetes fuera de la habitual interfaz BSD. Originalmente se pensó como sustituto de las herramientas que se habían venido utilizando para la implementación de *firewalls* (cortafuegos), denominadas *ipchains*, ya que su capacidad se había quedado limitada. La gran flexibilidad de la que se dotó a este entorno, ha favorecido la aparición de multitud de desarrollos basados en el mismo, desde *firewalls* para los que estaba pensado hasta *Network Address Translators* (NAT). Entre la multitud de implementaciones que se basan en *Netfilter* destacan los protocolos de enrutamiento para redes *ad hoc*.
Lo que proporciona es, básicamente, un conjunto de *hooks* o ganchos en los que el programador puede tomar el control sobre los paquetes durante su curso a lo largo de la pila de protocolos. Cuando un módulo basado en *Netfilter* se carga, se registran algunos de estos *hooks*. Cada vez que el programador toma el control sobre el paquete, y tras procesarlo según sus necesidades, puede tomar cinco posibles acciones:

- `NF_ACCEPT`: el paquete continuará con su recorrido normal.
- `NF_DROP`: el paquete será descartado (el propio Sistema Operativo se encarga de ello).
- `NF_STOLEN`: el usuario se encargará del paquete; en este sentido, el Sistema Operativo se 'desentiende' del mismo y será el propio código del usuario quien se responsabilice de su procesamiento (incluso de su eliminación).
- `NF_QUEUE`: el paquete se encola para su posterior manejo desde el espacio de usuario, tras ser recogido (por el driver `ip_queue`); estos paquetes se manejan de manera asíncrona.
- `NF_REPEAT`: en este caso se vuelve a llamar al mismo *hook*.

En concreto, la Figura 1 muestra la arquitectura de *Netfilter* en el caso de la capa IP, para su versión IPv4.

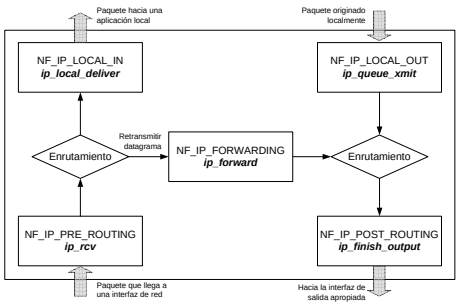


Figura 1: Arquitectura Netfilter IPv4

- **Tráfico Ascendente.** Al llegar una trama a través de una interfaz de red, esta llega al manejador correspondiente al protocolo IP. Tras pasar una serie de chequeos previos (longitud mínima, versión correcta del protocolo, checksum correcto,...) el datagrama llega al primer *hook* de *Netfilter*: `NF_IP_PRE_ROUTING`. Posteriormente, el datagrama atraviesa los procedimientos de enrutamiento IP y se pueden dar dos posibilidades: (1) el paquete es dirigido a la propia máquina, con la posibilidad de activar el gancho `NF_IP_LOCAL_IN`, o (2) el paquete tiene que ser re-enviado a través de una interfaz de salida (en este caso se activaría el gancho `NF_IP_FORWARDING`). En el segundo de los casos, el programador podría volver a tener el acceso al datagrama, a través del gancho `NF_IP_POST_ROUTING`, que se localiza en la última función que ejecutaría el protocolo IP.
- **Tráfico Descendente.** En este caso, la llegada del datagrama IP es desde las capas superiores (normalmente TCP o UDP), momento en el que el programador puede obtener el control sobre el mismo, utilizando el gancho `NF_LOCAL_OUT`. Si bien en la figura parece que los procedimientos de enrutamiento se ejecutan a continuación, la realidad es que ocurre al revés. En este punto, ambos sentidos (ascendente y descendente) convergen, por lo que se vuelve a tener acceso al datagrama a través del gancho `NF_IP_POST_ROUTING`.

3.2. Estructura de la implementación del DSR

La Figura 2 muestra la arquitectura de alto nivel de la implementación que se ha llevado a cabo, que consta de los siguientes elementos [5]:

- **Módulo principal o *core*.** El *core* o núcleo de la implementación interacciona con la implementación de la capa de protocolos que se emplea en el Sistema Operativo Linux y, por tanto, proporciona el punto de entrada al módulo, tanto para tráfico generado por el propio dispositivo (proveniente de las capas superiores) como aquel que llegue al interfaz de red. Además, se encarga de registrar

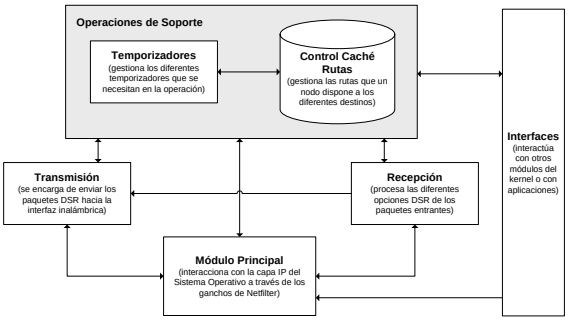


Figura 2: Estructura de la implementación del DSR

el módulo cuando este es cargado, habilitando las interfaces tanto con el espacio de usuario como con el resto de módulos del *kernel*.
Para capturar el tráfico, tiene dos rutinas principales, cada una de ellas asociadas a un gancho *Netfilter* diferente (`NF_IP_PRE_ROUTING` y `NF_IP_LOCAL_OUT` para los tráficos entrante y saliente, respectivamente). Además, en la entidad que permite la interconexión de una isla DSR con el exterior, *Gateway*, también se empleará el gancho `NF_IP_FORWARDING` (ver Figura 5).

- **Módulo de transmisión.** En este módulo se incluyen todas aquellas funcionalidades que se requieren para transmitir paquetes DSR, ya sea tráfico de control (propio de DSR) o transportando información del usuario.

- **Módulo de recepción.** Procesa los paquetes entrantes, en función de la opción DSR correspondiente.

- **Módulo de control de la caché de rutas.** Unos de los elementos principales del protocolo DSR es la estructura en la que mantine las rutas a los potenciales destinos de las comunicaciones. La implementación creada añade un módulo que se encarga de gestionar dicha estructura.

- **Módulo de temporizadores.** El funcionamiento del protocolo DSR requiere de la utilización de diversos temporizadores, que se gestionan en un mismo módulo de la implementación.

- **Interfaces.** Básicamente implementa las dos interfaces de comunicación del protocolo: con el espacio de usuario y con otros módulos del *kernel*.

Como se puede observar, los dos procedimientos básicos (descubrimiento y mantenimiento de ruta) de DSR se mapean sobre diversas funcionalidades de los módulos en los que se ha estructurado la implementación. La descripción se hará en base a esta estructura, ya que se corresponde mejor con la visión más tradicional del protocolo.

3.3. Incorporación del módulo DSR en la capa de protocolos IP

Al cargar el módulo DSR, se registran los puntos *'ganchos'* en los que el programador puede adquirir el control sobre el datagrama IP. Más concretamente, se distinguen tres puntos, en función de que el paquete en cuestión sea tráfico ascendente (caso en el que se emplearía el gancho `NF_IP_PRE_ROUTING`) o descendente, en el que se activaría el `NF_IP_LOCAL_OUT`). Además de la manera de capturar el tráfico, es importante analizar el modo en el que este es procesado a continuación, ya que de alguna manera se tiene que modificar el procesamiento 'tradicional' que la capa IP hace de cada uno de los datagramas, teniendo en cuenta que, fuera del módulo, el formato DSR no es entendido por ninguna entidad de la capa de protocolos.

3.3.1. Tráfico Descendente

El caso que hay que diferenciar en este punto es en el que un datagrama IP tiene que ser transmitido al siguiente salto en una ruta determinada, y este no coincide con el destino final del mismo. Debido a sus características intrínsecas de enrutamiento fuente, DSR no puede emplear la estructura de enrutamiento que maneja el Sistema Operativo, ya que esta se basa en la estructura *Destino_Final - Siguiente_Salto*. Además, desde el punto de vista de cada paquete, éste se tiene que transmitir a un nodo cuya dirección no coincide con el destino final, lo que implica ciertas dificultades desde el punto de vista del funcionamiento tradicional de la capa IP, por lo que se tiene que 'trucar' la interacción del protocolo IP con las capas inferiores, en particular con la capa MAC y el procedimiento ARP (*Address Resolution Protocol*). Otra particularidad que hay que considerar viene dada por la limitación impuesta por las interfaces físicas en cuanto al tamaño máximo de las tramas que se pueden transmitir; el módulo DSR se introduce en la capa de protocolos a través del marco proporcionado por *Netfilter*, pero el resto de las entidades no son conscientes de la sobrecarga que va a introducir, por lo que se tiene que asegurar, de alguna manera externa, que el tamaño total de la trama, incluyendo la sobrecarga introducida por DSR no supere el máximo soportado por el controlador correspondiente a la interfaz física.

3.3.2. Tráfico Ascendente

Se emplea `NF_IP_PRE_ROUTING` para capturar el paquete y se pueden dar dos posibilidades: el paquete es para un proceso (aplicación) local, con lo que el módulo DSR deberá extraer toda la información propia del datagrama, y dejar que siga su procesamiento habitual (`NF_ACCEPT`); por el contrario, se trata de un nodo intermedio o es un paquete de control DSR, casos en los que habrá que procesar la información DSR que lleve el datagrama, no dejando que siga su proceso habitual por parte de la capa IP.

3.3.3. Funcionalidad Gateway

Cuando uno de los equipos de la isla DSR cumpla la funcionalidad de *Gateway*, interconectándola con redes externas, también se tiene que habilitar el gancho `NF_IP_FORWARDING`, para poder procesar aquellos paquetes que provengan de una red externa y estén dirigidos a uno de los dispositivos que estén presentes en la red DSR.

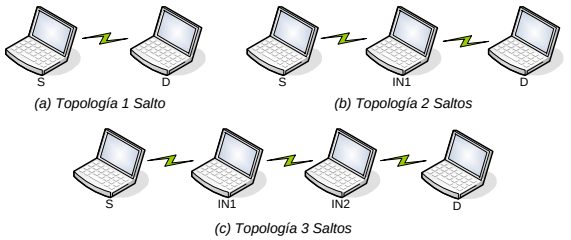


Figura 3: Topologías empleadas en la validación del DSR

4. Validación de DSR

4.1. Establecimiento de la plataforma

Una de las principales dificultades que aparecen a la hora de validar un protocolo de enrutamiento multi-salto es que las tecnologías de comunicaciones existentes en la actualidad no fueron diseñadas para este tipo de topologías. En particular IEEE 802.11b, que es la tecnología WLAN más extendida en la actualidad, define dos modos de funcionamiento: infraestructura y *ad hoc*. De ellas, la segunda es la que parece más apropiada para el establecimiento de configuraciones multi-salto, ya que la idiosincrasia claramente centralizada de la primera se contrapone a las mismas. Sin embargo, en la especificación del estándar 802.11 [6] se dice claramente que este modo de operación sólo es posible siempre y cuando *las estaciones puedan comunicarse entre sí directamente*, o lo que es lo mismo, pertenezcan al mismo área de cobertura. Si las estaciones formaran una topología multi-salto real, en la que dos nodos no contiguos no pertenecieran al mismo área de cobertura, el procedimiento de generación distribuida de *beacons* empleado por el estándar 802.11 no funcionaría correctamente y daría lugar a inestabilidades. Para establecer una plataforma estable de medidas se empleó la funcionalidad *mackill*, que se basa en el filtrado de las tramas provenientes de ciertas estaciones en función de su dirección MAC. Se podría decir que de esta manera no se plasma de manera totalmente fidedigna un escenario real, ya que aunque las estaciones estuvieran desconectadas *'virtualmente'* sí que estarían dentro del mismo área de cobertura y, por tanto, las transmisiones de las unas tendrían influencia en el rendimiento de las otras. Por otra parte, el problema del terminal oculto, que se da cuando dos estaciones que no están dentro del mismo área de cobertura transmiten de manera simultánea hacia una tercera estación, que se encuentra entre las dos anteriores, no podría ser analizado de esta manera. Ambas aseveraciones son teóricamente ciertas, aunque de manera práctica, hay que decir que los dispositivos inalámbricos comerciales disponibles se caracterizan por tener diferentes rangos de transmisión (en función de la tasa binaria de trabajo). Además el rango de detección de portadora (fundamental para el protocolo de acceso al medio que se está empleando) es claramente superior a todos ellos, de tal manera que la existencia '*real*' del problema del terminal oculto pueda no ser tan clara.

4.2. Sobrecarga del protocolo DSR

Un paso previo al análisis extenso de las prestaciones del protocolo de enrutamiento, es verificar que la sobrecarga que introduce, bien por los mecanismos de control que necesita, bien por las cabeceras necesarias en cada paquete, no es excesiva. Para ello se dispuso de tres configuraciones, de uno, dos y tres saltos (siempre asumiendo canal ideal), como muestra la Figura 3. Primeramente se realizará una configuración estática, en la que el enrutamiento se configura manualmente, modificando las tablas de enrutamiento de cada uno de los terminales. Posteriormente se establecerán las mismas topologías, utilizando el protocolo DSR, habilitando la desconexión virtual de los terminales mediante el uso de *mackill*.

Como se dijo anteriormente, se pueden emplear hasta tres mecanismos de reconocimiento diferentes en el procedimiento de mantenimiento de ruta; aunque desde el punto de vista de rendimiento, la solución óptima sería emplear la información proporcionada por la capa subyacente; en ese sentido, IEEE 802.11b utiliza un mecanismo de reconocimientos de nivel 2, lo que podría llevar a pensar (teniendo, en cuenta que será esta la tecnología que se empleará para validar el desarrollo realizado) que se debería emplear esta capacidad para acometer el procedimiento de mantenimiento de ruta. Lamentablemente las interfaces disponibles en la actualidad no exportan la información requerida, por lo que se tenían que explorar las otras alternativas. La posibilidad de los *reconocimientos pasivos*, además de que implica un gasto energético mayor, no se puede aplicar, debido a una limitación intrínseca del marco NETFILTER, que no permite procesar paquetes, recibidos en modo promiscuo, por parte del programador, pues los filtra el protocolo IP antes de que lleguen al primer *hook* que puede emplear. De esta manera, se tiene que incorporar el procedimiento de reconocimientos propios de DSR lo que, además, favorece una implementación más flexible. Por lo tanto, se tiene que demostrar que su efecto no es relevante.

Como se muestra en la Figura 4, el protocolo DSR no introduce una sobrecarga apreciable, independientemente que se use los reconocimientos propios de DSR o no (en este caso se asume que sí que existe la posibilidad de acceder a la información correspondiente en las capas inferiores). Los rendimientos instantáneos obtenidos en transmisiones de 60 segundos de duración han sido similares en los tres casos analizados,

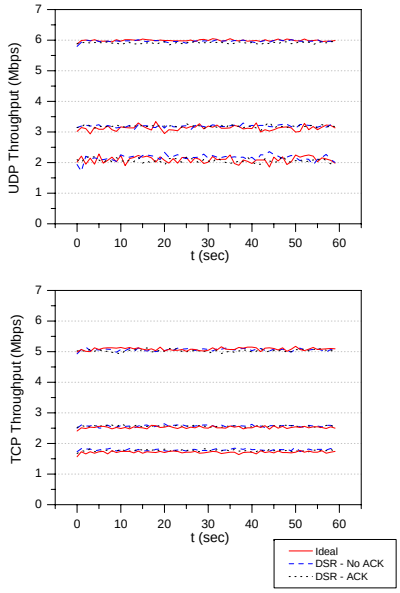


Figura 4: Throughput instantáneo

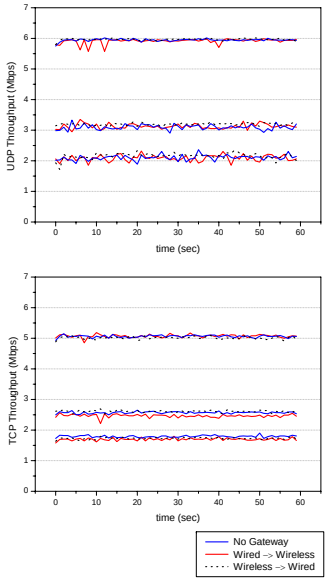


Figura 6: Throughput instantáneo con *Gateway*

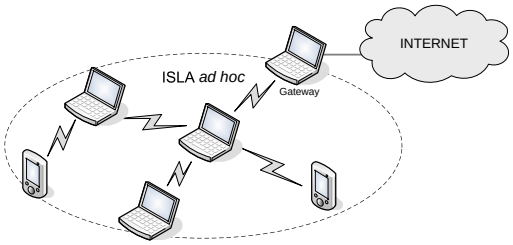


Figura 5: Funcionalidad Gateway en redes multi-salto

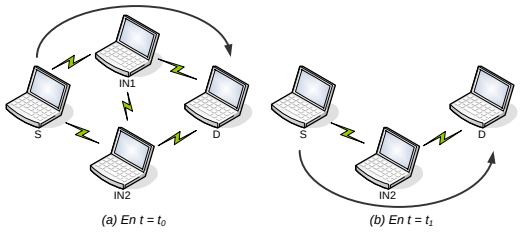


Figura 7: Topología para el cambio de ruta

para uno, dos y tres saltos. El resultado se cumple tanto con tráfico UDP como con TCP.

Una de los aspectos más innovadores de la implementación que se ha realizado es que permite la interconexión de *islas ad hoc*, basadas en el protocolo DSR con otras redes, a través de una entidad *Gateway*, como se muestra en la Figura 5; es, por tanto, interesante comprobar que su presencia no añade una sobrecarga inapropiada. Las topologías que se muestran en la Figura 3 siguen siendo válidas, aunque ahora el nodo que antes hacía las veces de destino se le asigna el rol de *Gateway*, con lo que incorpora una interfaz *Ethernet* a través de la cual se conecta con un nodo destino fuera de la *isla ad hoc* (*correspondent node* en la terminología de *Mobile IP*). Como se muestra en la Figura 6, su presencia no es apreciable desde el punto de vista del rendimiento, ya que independientemente de que el tráfico vaya desde o hacia la *isla ad hoc* y del número de saltos dentro de esta, no se observa disminución en el caudal medido.

4.3. Dinamismo del DSR

Una de las ventajas que introduce un protocolo de enrutamiento para redes *ad hoc* frente a una configuración manual (como la que se ha usado en la Sección anterior para validar los resultados) es que este es capaz de adaptarse a los cambios topológicos de la red; como se explicó anteriormente, tras detectar la caída en algún enlace, un nodo avisa a la fuente de la ruta correspondiente, para que trate de localizar una ruta alternativa. Para poner de manifiesto este hecho se utilizó una topología en rombo, como se ve en la Figura 7; el nodo S transmite cierta información al nodo D. Como no se puede establecer una comunicación directa entre ellos, tienen que utilizar a un tercer nodo ($IN1$) para que retransmita el tráfico entre ambos. En un momento de la comunicación, este último terminal desaparece, y la fuente de la comunicación que encuentra una ruta alternativa a través del nodo $IN2$.

La Figura 8 muestra el rendimiento instantáneo obtenido sobre la topología descrita anteriormente, con tráfico tanto UDP como TCP. Se pueden extraer dos conclusiones principales:

- Se demuestra la correcta adaptación del protocolo de enrutamiento a cambios topológicos. En la Figura 8 destaca claramente el momento del cambio de ruta, con una caída brusca del rendimiento, aunque éste se recupera de manera casi inmediata, correspondiéndose con el instante en el que la fuente de la comunicación encuentra una ruta alternativa hacia el destino.
- También es importante destacar las diferencias existentes entre los protocolos TCP y UDP. Como

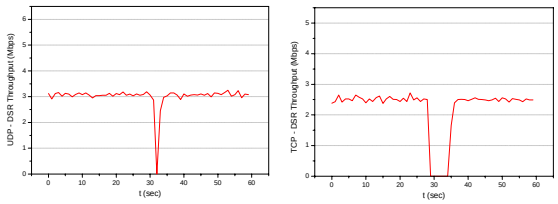


Figura 8: Throughput Instantáneo en la Topología 1

se puede observar en la Figura 8, el tiempo en el que el rendimiento instantáneo se anula es sensiblemente superior en el caso de TCP. Cuando la entidad TCP transmisora detecta errores en la transmisión, aplica sus mecanismos de control de flujo (pues asume que los errores se deben a la saturación de elementos intermedios de la red) y, por tanto, reduce la tasa de transmisión; sin embargo, como protocolo de transporte orientado a la conexión, asegura la correcta recepción, por parte del receptor TCP, de todos los segmentos. Por el contrario, en el caso de UDP, el nodo S sigue con la transmisión de información continuamente, con lo que se incurre en ciertas pérdidas, que se sitúan en torno al $2\,\%$ en esta situación en particular.

4.4. Comportamiento de los protocolos IP sobre redes multi-salto

Una vez que se ha demostrado que la sobrecarga introducida por el protocolo DSR implementado puede considerarse despreciable, éste se puede emplear para caracterizar el comportamiento de diferentes protocolos sobre topologías multi-salto, utilizando IEEE 802.11b como tecnología subyacente. El análisis se centrará en la pila de protocolos más extendida en la actualidad: IP. En primer lugar se estudiará el comportamiento de UDP. Se trata de un protocolo de transporte no orientado a la conexión, por lo que es el más idóneo a la hora de analizar la capacidad de una red, ya que no incorpora ningún mecanismo de control y su sobrecarga es muy reducida. Las topologías que se presentaron anteriormente (ver Figura 3) son las que se emplearon para realizar estas medidas, sobre las que se realizaron diez experimentos independientes (con transmisiones de 30 segundos de duración); la Figura 9 muestra los resultados obtenidos.
Una primera conclusión que se puede extraer es la estabilidad de la implementación realizada, ya que los experimentos individuales arrojan rendimientos muy similares (con muy poca variabilidad entre ellos); además, los resultados para un único salto son idénticos a los que ya se obtuvieron en [7], lo que sigue poniendo de manifiesto que el protocolo DSR no introduce ninguna sobrecarga. Por otra parte, y como se puede observar en la figura, el rendimiento sigue una ley decreciente con el número de saltos (N), de manera que los valores obtenidos para dos y tres saltos se corresponden, aproximadamente, con el obtenido para un salto

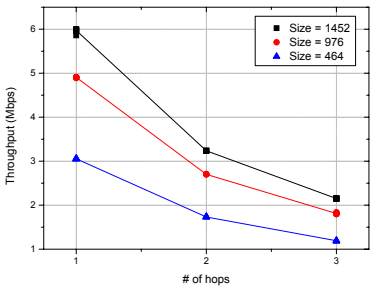


Figura 9: Rendimiento UDP en redes multi-salto

dividido entre N. De alguna manera, este comportamiento, puede poner un límite práctico en el número máximo de saltos que se pueden emplear en situaciones reales, ya que la disminución del rendimiento es apreciable a medida que N crece. Además, como parecía lógico, a medida que se disminuye la carga útil transmitida, el rendimiento obtenido decrece paulatinamente, ya que la sobrecarga introducida es independiente de los datos transmitidos y su relevancia es, por tanto, mayor al disminuir el tamaño de los datos [7].
Una vez analizado el comportamiento del protocolo UDP se estudiará el que presenta TCP. La mayoría de las aplicaciones Internet empleadas en la actualidad (`web, email, ftp`) se basan en este protocolo de transporte que, al contrario que el anterior, ofrece un servicio orientado a la conexión, con complejos mecanismos de control de errores y de flujo. En este caso, además de modificar la longitud de la carga útil transmitida, también se variará el tamaño de la ventana TCP, que marca la cantidad de segmentos que el transmisor puede enviar sin recibir confirmación por parte del receptor. Se realizaron cinco medidas independientes, para cada tamaño de segmento y ventana.
La Figura 10 muestra los resultados obtenidos. Como se puede observar, el rendimiento obtenido crece con la ventana TCP empleada, aunque para valores mayores de cinco segmentos, el aumento es prácticamente inapreciable, por lo que se puede concluir que utilizar ventanas mayores de 5 segmentos no aporta beneficio alguno. Como sucedía con UDP, los valores de *throughput* disminuyen con la longitud útil (el *Maximum Segment Size* o MSS en este caso). Por otra parte, que el rendimiento sigue la misma tendencia que con el protocolo UDP, ya que se va dividiendo paulatinamente por el número de saltos, para las mismas condiciones de MSS y ventana TCP. Los resultados obtenidos para el caso de un salto coinciden con los reportados en [8, 9].

5. Conclusiones y líneas futuras de investigación

La actividad en el área de las redes inalámbricas multi-salto está atrayendo una extensa labor investigadora recientemente; sin embargo, como ya se ha dicho previamente, no se dispone aún de un conocimiento en lo que se refiere al comportamiento sobre arquitecturas reales

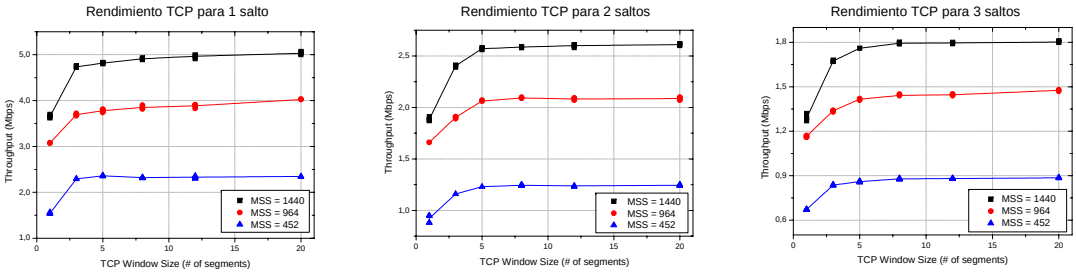


Figura 10: Rendimiento TCP en redes multi-salto

de este tipo de protocolos, lo que es un aspecto fundamentar para verificar su escalabilidad e idoneidad. Es en este punto donde se sitúa el aspecto más innovador de este trabajo, ya que acomete la implementación completa de un algoritmo de enrutamiento para redes multi-salto, y su comportamiento se estudia desde un punto de vista completamente experimental. Es importante mencionar que el protocolo DSR que se ha implementado no introduce una sobrecarga relevante, incluso cuando se utiliza uno de los dispositivos como *Gateway*, conectándole a redes externas.

Aprovechando la implementación realizada y una vez comprobado que no introduce sobrecarga adicional, se utiliza para acometer una caracterización del comportamiento que los protocolos tradicionales IP presentan sobre redes inalámbricas multi-salto, siguiendo un enfoque completamente experimental. Una primera conclusión que se puede extraer es que, con la tecnología que hay disponible en la actualidad, el uso que se puede hacer de topologías multi-salto es bastante limitado; será importante tratar de utilizar este tipo de protocolos en entornos multi-radio, en los que la utilización de un salto (o enlace) no implicaría interferencia en las otras, aunque todavía no se dispone de protocolos (de más bajo nivel) que puedan gestionar dichas capacidades.

En cuando a las futuras líneas de trabajo que se abren, se tratará de aprovechar las capacidades que tiene el módulo que se ha creado para seguir con el análisis que se ha empezado a realizar. Además, aprovechando el conjunto de interfaces que se han incorporado, y el enrutamiento fuente que DSR utiliza de manera intrínseca, se tratará de ampliar el alcance de su funcionamiento, tratando de utilizar métricas diferentes para elegir las rutas, ya que como se recoge en [10], el *camino más corto no siempre es el óptimo*.

Referencias

[1] L. Muñoz, R. Agüero, J. Choque, J. A. Irastorza, L. Sánchez, M. Petrova y P. Mähönen. Empowering next-generation wireless personal communication networks. *IEEE Communications Magazine*, 42(5):64–70, mayo 2004.

[2] M. Berg, J. Markendahl, I. Herwono, S. Göbbels, R. Pabst, R. Agüero, J. Choque, L. Muñoz, M. Prytz y O. Strandberg. Non-conventional / low-cost concepts. Deliverable D2.3, Ambient Networks (IST-2002-507134), junio 2004.

[3] D. B. Johnson, D. A. Maltz y Y. C. Hu. The dynamic source routing protocol for mobile ad hoc networks (DSR). Internet Draft Version 10, IETF, julio 2004.

[4] R. Russell y H. Welte. Linux netfilter hacking HOWTO. Linux Howto, julio 2002.

[5] V. Gutiérrez, R. Agüero y L. Muñoz. Routing mechanism for ad-hoc WPAN environment. Deliverable D5.4, Power Aware Communications for Wireless OptiMised personal Area Network - PACWOMAN (IST-2001-34157), septiembre 2004.

[6] IEEE Computer Society LAN MAN Standards Committee. *IEEE 802.11: Wireless LAN Medium Access Control and Physical Layer Specifications*, agosto 1999.

[7] R. Agüero, M. García, J. Choque y L. Muñoz. Caracterización del protocolo UDP sobre redes de Área local inalámbricas IEEE 802.11b. En *Actas de las III Jornadas de Ingeniería Telemática, JITEL*, páginas 235–241. septiembre 2001.

[8] M. García, R. Agüero y L. Muñoz. On the unsuitability of TCP RTO estimation over bursty error channels. En *Proceedings of the IFIP TC6 9th International Conference on Personal Wireless Communications, PWC - LCNS 3260*, páginas 343–348. septiembre 2004.

[9] R. Agüero, L. Sánchez, M. García, J. Choque y L. Muñoz. Análisis experimental del comportamiento de TCP sobre IEEE 802.11b y del protocolo snoop como mecanismo de mejora. En *Actas de las IV Jornadas de Ingeniería Telemática, JITEL*, páginas 205–212. septiembre 2003.

[10] D. D. Couto, D. Aguayo, B. Chambers y R. Morris. Performance of multihop wireless networks: Shortest path is not enough. En *Proceedings of the First Workshop on Hot Topics in Networking*. octubre 2002.

Modelado de una red IEEE 802.11 con tráfico heterogéneo

B.Bellalta[†], M.Meo[‡], A.Escudero[†], M.Oliver[†]

[†]Dept. de Tecnologia. Universitat Pompeu Fabra
Passeig de la Circumval.lació 8.
08003 - Barcelona (Barcelona)
Teléfono: +34-93-5422945 Fax: +34-93-5422517
E-mail: boris.bellalta@upf.edu

[‡]Dipartimento di Elettronica. Politecnico di Torino
Corso Duca degli Abruzzi 24.
10129 - Torino (Italy)
Teléfono: +39-011-5644053 Fax: +39-011-564099
E-mail: michela@polito.it

Abstract *This work presents a model of the MAC layer which characterizes the performance of an IEEE 802.11 network using traffic under non-saturation conditions. Main results evaluate the effect of simultaneous multiplexing of different types of traffic flows over a finite and shared set of resources and how are their interactions and dependencies among them. Once the model is described, it is validated using results of NS-2 simulations. Furthermore, we present a set of different scenarios where our model could be applied. Finally, a real infrastructure scenario is presented with coexistence of elastic TCP (FTP) flows in the downlink and bidirectional rigid UDP traffic.*

1. Introducción

El tráfico de información en Internet está compuesto principalmente por flujos de tráfico elástico generados por la transferencia de objetos de datos (documentos, correos electrónicos, páginas web, canciones, imágenes,... etc). Estos flujos utilizan a nivel de transporte el protocolo TCP y su principal característica es que adaptan la tasa (el ancho de banda que consumen) al estado de la red. A causa del aumento de la capacidad en las redes de acceso, están empezando a ser utilizados nuevos servicios y aplicaciones que permiten transmitir voz y vídeo en tiempo real sobre Internet. La principal característica de estos flujos, que llamaremos rígidos, será la necesidad de disponer de un ancho de banda determinado (de ahí su nombre) para su correcto funcionamiento.

En este contexto, las redes WLAN basadas en el protocolo IEEE 802.11 [1], han de gestionar tanto tráfico de características elásticas como de características rígidas. El comportamiento del protocolo IEEE 802.11 ha sido ampliamente estudiado en múltiples configuraciones, destacar el trabajo de Bianchi [2], que ha servido de punto de referencia en la mayoría de estudios posteriores. En [2, 3] se propone un modelo para evaluar las prestaciones del protocolo MAC IEEE 802.11 en condiciones de saturación (todos los nodos siempre tienen un paquete preparado para ser transmitido), obteniendo una solución elegante a lo que se puede considerar el límite superior de las prestaciones de la red. Los resultados obtenidos considerando la red en estado de saturación distan de la realidad (una red nunca trabajará correctamente en estas condiciones) y no tienen una aplicabilidad real. Así para el estudio de su capacidad, planificación de red, diseño y evaluación de mecanismos de gestión de recursos, etc., se necesita un modelo que también incluya las características de tráfico de cada nodo en condiciones de no saturación (un nodo no está continuamente transmitiendo) [5]. Entre éstos, en [4] se propone una simple modificación del modelo propuesto por Bianchi, pero que permite modelar con suficiente precisión las condiciones de no saturación. En cualquier caso, en estos trabajos no se estudian las relaciones entre flujos de tráfico de diferentes características, encontrándose como uno de los primeros trabajos el descrito en [6].

2. Descripción de los flujos de tráfico

Es difícil realizar una clasificación genérica de los diferentes flujos de tráfico que existen en Internet, la más simple consiste en considerar dos tipos de flujos: elásticos y rígidos, en función del protocolo de transporte utilizado, TCP (Transmission Control Protocol) y UDP (User Datagram Protocol). A continuación se detallan las características consideradas para ambos tipos de tráfico.

2.1. Tráfico rígido (UDP)

Un flujo de tráfico rígido quedará caracterizado por un ancho de banda que permanecerá constante mientras el flujo esté activo. Si la red puede asegurar el ancho de banda requerido por el flujo, se desarrollará correctamente. En caso contrario, normalmente el flujo es finalizado. Para caracterizar este tipo de flujo es necesario conocer el ancho de banda requerido

[1]Este trabajo se ha realizado en el marco del proyecto CICYT TIC2003-09279-C02-01, Newcom (Network of Excellence in Wireless COMmunications) y i2CAT (Internet 2 a Catalunya).

B_s. Conociendo la longitud media L_s de los paquetes, se puede determinar la tasa de paquetes que ha de entregar a la red por segundo, $\lambda_s = B_s/L_s$. En el presente trabajo se considera la presencia de K classes de tráfico streaming, cada una con un ancho de banda B_s^k, una longitud media de los paquetes L_s^k y una tasa media de generación de paquetes por segundo igual a λ_s^k.

2.2. Tráfico elástico (TCP)

Caracterizar de manera correcta el tráfico TCP presenta una elevada dificultad a causa de la gran cantidad de parámetros con una dinámica compleja a considerar, que depende altamente de cada situación concreta [9].

En este trabajo se consideran únicamente flujos de tráfico elástico ideales. Para caracterizar estos flujos también se utiliza el ancho de banda que consumen, pero en este caso, el valor del ancho de banda utilizado por cada flujo dependerá del estado de la red. Así un enlace con un ancho de banda B es perfectamente compartido por n_e flujos de tráfico elástico si cada flujo utiliza un ancho de bada igual a $B_e = B/n_e$ [8]. De manera equivalente, el ancho de banda que utiliza un flujo elástico es $B_e = L_e/X(n_e)$, donde $X(n_e)$ es el tiempo necesario para transmitir un paquete de longitud L_e *bits* en presencia de n_e flujos elásticos. Para cada flujo elástico la tasa de generación de paquetes será $\lambda_e = 1/X(n_e)$ que equivale a la máxima velocidad de transmisión de paquetes μ_e para cada flujo. En estas condiciones, la utilización de la cola de transmisión será igual a $\rho_e = \lambda_e/\mu_e = 1$ (siempre habrá un paquete listo para ser transmitido) y por tanto el flujo, mientras esté activo, puede ser modelado asumiendo condiciones de saturación. La variación del ancho de banda en función del número de flujos activos implica obviamente una variación en el tiempo de respuesta del flujo (tiempo necesario para transmitir los datos asociados al flujo).

2.3. Tráfico rígido y elástico simultáneamente

En una situación ideal, la multiplexación de tráfico rígido y elástico sobre el mismo ancho de banda B conduce a la siguiente situación [8]

$$B_e = \frac{B - \sum_k n_s^k B_s^k}{n_e} \tag{1}$$

si se cumple que $\sum_k n_s^k B_s^k \leq B$, en caso contrario $B_e = 0$. A partir de (1) se extrae que los flujos rígidos no se ven afectados por el tráfico elástico, cosa que aunque considerando flujos elásticos ideales no se cumplirá si se utiliza a nivel MAC un protocolo de acceso al medio bajo contienda. Lo que es cierto es que en cualquier caso, existe una dependencia entre la presencia de flujos rígidos y las prestaciones de los flujos de tráfico elástico.

3. Modelo

Por motivos de falta de espacio y como ya ha sido ampliamente descrito en múltiples artículos, no entraremos a detallar el funcionamiento del protocolo MAC IEEE 802.11 (remitimos a las referencias). A continuación presentamos el modelo, definiendo el tipo de nodo considerado y analizando los efectos del protocolo MAC tanto para una única clase de tráfico (todos los nodos iguales) como para un caso genérico con K clases de terminales móviles (nodos) con perfiles de tráfico diferentes.

3.1. Descripción de un nodo

Cada terminal móvil se ha modelado como una cola $M/M/1/Q$ donde Q es el tamaño de la cola en número de paquetes (considerando el paquete que está en servicio). El tráfico ofrecido a un nodo i cualquiera es

$$\nu_i(\mathbf{n}) = \lambda_i X_i(\mathbf{n}) \tag{2}$$

donde $X_i(\mathbf{n})$ es el tiempo de servicio de un paquete en el nodo i y λ_i es la tasa de llegada de paquetes a la cola de transmisión de nivel MAC del nodo i en presencia de $\mathbf{n}$ nodos con K perfiles de tráfico diferentes. En este contexto, $\mathbf{n}$ es un vector de K ($m = 1..K$) posiciones con $\mathbf{n}(k)$ el número de terminales móviles con las mismas características de tráfico. El tráfico ofrecido por cada nodo a la red es

$$\rho_i(\mathbf{n}) = \lambda_i(1 - P_{b,i}(\mathbf{n}))X_i(\mathbf{n}) \tag{3}$$

donde la probabilidad de bloqueo para una cola $M/M/1/Q$ se obtiene a partir de

$$P_{b,i}(\mathbf{n}) = \frac{(1 - \nu_i(\mathbf{n}))\nu_i(\mathbf{n})^Q}{(1 - \nu_i(\mathbf{n})^{Q+1})} \tag{4}$$

y la ocupación media de la cola

$$E\{N_i(\mathbf{n})\} = \frac{\nu_i(\mathbf{n})}{1 - \nu_i(\mathbf{n})} - \\ - \frac{Q}{(1 - \nu_i(\mathbf{n}))^{Q+1}}\nu_i(\mathbf{n})^{Q+1} \tag{5}$$

con el caso especial de $\nu_i(\mathbf{n}) = 1$ donde $E\{N_i(\mathbf{n})\} = Q/2$ y en consecuencia, el tiempo medio de estancia de un paquete en la cola $E\{R_i(\mathbf{n})\} = E\{N_i(\mathbf{n})\}/\lambda_i$.

Se asume que el tiempo de servicio de un paquete sigue una distribución exponencial (o de manera equivalente, los paquetes tienen una longitud distribuida exponencialmente) y que llegan al nivel MAC siguiendo una distribución de Poisson (con estas suposiciones podemos obtener un modelo simple y tratable analíticamente). Es importante remarcar que las prestaciones que observa un nodo dependen de las características de tráfico de los otros nodos ya que, al ser un protocolo de acceso aleatorio, el tiempo de servicio depende del estado de la red y, por tanto, los otros parámetros relativos a las prestaciones de cada nodo también dependerán de como se encuentre la red.

3.2. Una única clase de tráfico

Considerando que todos los nodos presentan las mismas características de tráfico, es decir $K = 1$, el vector $\mathbf{n}$ se reduce al entero n y, omitiendo el índice i, los parámetros que caracterizan el perfil de tráfico de los n nodos activos en la red son: un ancho de banda B y una longitud media de los paquetes L. La tasa de llegada de paquetes a nivel MAC se obtiene a partir de los parámetros anteriores, siendo $\lambda = B/L$ paquetes por segundo. Asumiendo que los nodos se comportan de manera independiente [2], la probabilidad que un nodo transmita un paquete es

$$\tau(n) = \frac{\varrho(n)}{E\{W(n)\}} \tag{6}$$

donde $\varrho(n)$, es la probabilidad que un nodo disponga de un paquete listo para ser transmitido. En general, esta probabilidad será la probabilidad que en la cola haya al menos un paquete, es decir, la utilización de la cola $\varrho(n) = \rho(n)$. $E\{W(n)\}$ es el número medio de ranuras que hace falta esperar para transmitir un paquete (uniformemente distribuidas). La expresión de la ecuación (6) se obtiene de una manera similar a la presentada en [4]. Llamando $P[NT]$ a la probabilidad que el nodo no transmita en una cierta ranura, la probabilidad de colisión p es

$$p(n) = 1 - P[NT]^{n-1} = 1 - (1 - \tau(n))^{n-1} \tag{7}$$

Existirá colisión si alguno de los $n-1$ nodos restantes transmite en la misma ranura. Considerando $P[QE]$ la probabilidad que la cola de transmisión de nivel MAC esté vacía y $P[QNE]$ la probabilidad que no esté vacía, $P[NT]$ se puede expresar como

$$\begin{aligned} P[NT] &= P[NT|QE] \cdot P[QE] + \\ &+ P[NT|QNE] \cdot P[QNE] = \\ &= 1 \cdot (1 - \rho(n)) + \\ &+ \frac{E\{W(n)\} - 1}{E\{W(n)\}} \cdot \rho(n) \end{aligned} \tag{8}$$

quedando reducida a

$$P[NT] = 1 - \frac{\rho(n)}{E\{W(n)\}} \tag{9}$$

de donde se desprende la expresión descrita en (6) si $\varrho(n) = \rho(n)$. En estas condiciones la probabilidad de colisión, des del punto de vista de un nodo, es la probabilidad de que como mínimo transmita otro de los $n-1$ nodos restantes en la misma ranura temporal

$$p(n) = 1 - (1 - \tau(n))^{n-1} = 1 - \left(1 - \frac{\varrho(n)}{E\{W(n)\}}\right)^{n-1} \tag{10}$$

A partir de la expresión anterior, se hace obvio que las prestaciones de la red, en este caso la probabilidad de colisión dependen de las características del tráfico de cada nodo. El tiempo de servicio observado por un nodo es

$$\begin{aligned} X(n) =& N_T(n) \left(\frac{E\{W(n)\}}{2}\alpha(n)\right) + \\ &+ (N_T(n) - 1)\, T_c(L) + T(L) \end{aligned} \tag{11}$$

donde $N_T(n)$ es el número de transmisiones necesarias para que un paquete sea recibido correctamente (no se ha considerado la presencia de errores en el canal, por tanto todas las retransmisiones son causadas por colisiones). La duración de una colisión es $T_c(L)$ y en caso de transmisión correcta es $T(L)$ donde L es la longitud del paquete a transmitir (incluyendo las diferentes cabeceras de nivel de enlace y físico). El parámetro $\alpha(n)$ corresponde a la duración media de una ranura, contemplando que el contador de ranuras se congela cuando el canal se observa utilizado, ya sea por una colisión o una transmisión libre de colisiones.

Considerando que un paquete no se deshecha una vez se ha llegado al número máximo de transmisiones

$$N_T(n) \approx \sum_{k=1}^{\infty} k(1-p)p^{k-1} = \left(1 - \frac{\varrho(n)}{E\{W(n)\}}\right)^{1-n} \tag{12}$$

Cuando un nodo en periodo de *back-off* detecta que el canal está ocupado, congela el contador de ranuras. El periodo temporal en que mantiene congelado el contador depende de la causa que hace que el canal esté ocupado: una transmisión correcta $T(L)$ o una colisión $T_c(L)$. En caso que el canal esté libre durante σ segundos, el contador se decrementa. Por tanto, el contador de ranuras será decrementado después de $T(L) + \sigma$ segundos con probabilidad p_s, $T_c(L) + \sigma$ segundos con probabilidad p_c o σ con probabilidad p_e.

La probabilidad que el canal este libre es la probabilidad que ninguno de los restantes $n-1$ nodos transmita en aquella ranura, por tanto

$$p_e(n) = \left(1 - \frac{\varrho(n)}{E\{W(n)\}}\right)^{n-1} \tag{13}$$

La probabilidad que una estación en backoff observe una transmisión correcta es la probabilidad que transmita únicamente uno de los restantes $n-1$ nodos

$$p_s(n) = (n-1)\tau(1 - \tau(n))^{n-2} \tag{14}$$

Finalmente, la probabilidad que una estación observe que en el canal se ha producido una colisión es $p_c(n) = 1 - p_e(n) - p_s(n)$. Considerando estas tres probabilidades, la duración media de una ranura es

$$\begin{aligned} \alpha(n) =& p_e(n)\sigma + p_s(n)(T(L) + \sigma) + \\ &+ p_c(n)(T_c(L) + \sigma) \end{aligned} \tag{15}$$

Obviamente, cuando en el sistema hay únicamente un nodo, $p_s(1) = 0$. Con dos nodos la probabilidad de observar una transmisión útil es la probabilidad que el otro terminal transmita, así pues: $p_s(2) = \tau(2)$. En consecuencia, el parámetro $\alpha(n)$ se incrementa con el

número de terminales móviles y por tanto la duración del periodo de espera aleatoria antes de transmitir se adapta al número de flujos activos.

El caudal del sistema (en *bps*) es

$$S(n) = n\frac{\rho(n)}{X(n)}L = n\lambda(1 - P_b(n)) \tag{16}$$

Finalmente, $E\{W(n)\}$ se obtiene a partir de los resultados presentados en [2] donde para una red con saturación se tiene

$$E\{W(n)\} = \\ = \frac{2(1-2p)}{(1-2p)(CW_{min}+1) + pCW_{min}(1-(2p)^m} \tag{17}$$

3.3. Generalización para K clases de tráfico

Una vez presentado el modelo considerando que todos los nodos presentan unos perfiles de tráfico homogéneos, la existencia de diferentes clases de tráfico hace necesario considerar las dependencias recíprocas entre los nodos. Como veremos estas dependencias quedan bien capturadas únicamente a partir de la utilización de la cola $\rho(n)$ ya que depende tanto de las características del tráfico generado por un nodo como del estado de la red. La probabilidad de transmisión de un nodo de tipo k es

$$\tau^k(\mathbf{n}) = \frac{\rho^k(\mathbf{n})}{E\{W^k(\mathbf{n})\}} \tag{18}$$

De la misma manera, las otras expresiones también quedan igualmente modificadas. Para el cálculo de la probabilidad de colisión hace falta considerar todos los posibles productos cruzados entre diferentes clases de tráfico (múltiples colisiones). Llamando $\theta(\mathbf{n})$ al conjunto de múltiples colisiones, la probabilidad de colisión observada por un nodo en presencia de K clases de nodos con tráfico diferente es

$$p^k(\mathbf{n}) = (1 - (1 - \tau^k(\mathbf{n}))^{\mathbf{n}(k)-1}) + \\ + \sum_{\bigvee j,j\neq k}^{K} (1 - (1 - \tau^j(\mathbf{n}))^{\mathbf{n}(j)}) - \theta_s(\mathbf{n}) \tag{19}$$

Un resultado interesante que se puede extraer de la expresión anterior es que nodos con flujos de tráfico reducido observan una probabilidad de colisión superior a flujos de tráfico de mayor ancho de banda.

Para el cálculo de la duración media de una ranura, hace falta hacer las mismas consideraciones anteriormente descritas para obtener p_e^k, p_s^k y p_c^k. Así, para el cálculo de p_e^k.

$$p_e^k(\mathbf{n}) = ((1 - \tau^k(\mathbf{n})^{\mathbf{n}(k)-1}) \cdot \\ \cdot \prod_{\bigvee j,j\neq k}^{K} (1 - \tau^j(\mathbf{n}))^{\mathbf{n}(j)}) \tag{20}$$

Para el cálculo de $p_s^k(\mathbf{n})$ se ha de considerar la probabilidad que únicamente transmita un único nodo de entre las K clases de nodos. Si $p_{s,k}(\mathbf{n})$ es la probabilidad que únicamente transmita un nodo de la clase k, tendremos

$$p_s^k(\mathbf{n}) = \sum_{\bigvee u} p_{s,u}(\mathbf{n}) \prod_{\bigvee j,j\neq u}^{K} (1 - \tau^j(\mathbf{n}))^{\mathbf{n}(j)} \tag{21}$$

En el caso $k = u$ hace falta considerar únicamente que un nodo entre los $n-1$ nodos restantes transmita. Finalmente, $p_c^k(\mathbf{n}) = 1 - p_e^k(\mathbf{n}) - p_s^k(\mathbf{n})$.

Con K clases de tráfico, hace falta también considerar la presencia de paquetes con diferentes longitudes, es decir, podemos tener transmisiones correctas y colisiones de diferentes duraciones. En esta situación la duración media de una ranura para un nodo de la clase k es

$$\alpha^k(\mathbf{n}) = p_e^k(\mathbf{n})\sigma + \\ + p_s^k(\mathbf{n}) \left(\sum_{\bigvee j} \beta^j (T(L^j) + \sigma) \right) + \\ + p_c^k(\mathbf{n}) \left(\sum_{\bigvee j} \beta^j (T_c(L^j)) + \sigma) \right) \tag{22}$$

con β^j la probabilidad que un nodo con tráfico de tipo j transmita un paquete. Para obtener esta probabilidad, ponderamos el número de nodos de cada tipo con su probabilidad de transmisión

$$\beta^k = \frac{\mathbf{n}(k)\tau^k(\mathbf{n})}{\sum_{\bigvee j} \mathbf{n}(j)\tau^j(\mathbf{n})} \tag{23}$$

Finalmente, una métrica de interés, para entender el funcionamiento de la red es el coeficiente medio de utilización, que nos indica cuando la red entra en saturación

$$\rho_{red} = \frac{\sum_k \mathbf{n}(k)\rho^k}{\sum_k \mathbf{n}(k)} \tag{24}$$

La estructura de las tramas no se adjunta y se puede consultar en [1].

3.4. Validación y Resultados

A causa de la dependencia de las expresiones con la utilización de la cola $\rho^k(\mathbf{n})$ y el conjunto de ecuaciones no lineales que forman (6) y (10), no es posible obtener una solución cerrada al modelo presentado. Para su resolución se han utilizado técnicas iterativas de análisis numérico. La validación se ha realizado comparando los resultados obtenidos utilizando el modelo con resultados obtenidos por simulación. Con esta finalidad se ha programado, con la máxima rigurosidad, un simulador del protocolo MAC IEEE 802.11 utilizando las librerías COST (Component Oriented Simulation Toolkit) [13]. Los valores de los parámetros utilizados se muestran en la Tabla 1.

Parameter	Value
SIFS	10 μs
DIFS	50 μs
EIFS	364 μs
σ	20 μs
(CW_{min}, CW_{max})	$(32, 1024)$
MAC header	224 bits
PHY header	192 bits
L_{data}	MAC h.+L+PHY h.
L_{ack}	112+PHY header bits
RTS	160+PHY header bits
CTS	112+PHY header bits
R_{data}	2 Mbps
R_{basic}	2 Mbps

Tabla 1: Parámetros MAC IEEE 802.11

Los parámetros que caracterizan las dos clases de tráfico (nodos) consideradas son: tráfico rígido con un ancho de banda $B_s = 64$ $Kbps$ y $L_s = 1024$ $bits$ y tráfico elástico con un ancho de banda variable B_e y $L_e = 4096$ $bits$. La longitud de las colas en cada nodo está fijada a un valor de $Q = 50$ paquetes (incluido el paquete en servicio). Los resultados obtenidos por análisis se muestran prácticamente idénticos a los obtenidos por simulación, tanto en los valores numéricos como en las tendencias.

En la Figura 1 se muestra el caudal agregado únicamente para los flujos de tráfico rígido de 64 $Kbps$ cuando en la red hay presentes también n_e nodos con tráfico elástico, considerándose la utilización del mecanismo de acceso básico. El caudal se representa en función del número de flujos rígidos presentes en la red. Así, cuando en la red únicamente hay flujos rígidos, el caudal máximo se obtiene para $n_s = 12$ flujos, ya que a partir de este número de flujos la red entra en saturación y la utilización de la cola se mantiene constante e igual a la unidad ($\rho = 1$).

La inclusión en la red de flujos elásticos supone una reducción drástica en el número máximo de flujos rígidos (aspecto que contradice el comportamiento ideal de los flujos elásticos [8]). Como podemos observar, cuando en la red están presentes dos flujos de tráfico elástico, $n_e = 2$, el número máximo de flujos rígidos se limita a $n_s = 8$ y se reduce hasta únicamente $n_s = 2$ cuando el número de flujos elásticos aumenta hasta $n_e = 5$. La justificación por el comportamiento no ideal de los flujos elásticos está motivada por el propio funcionamiento del protocolo de enlace ya que intenta que todos los nodos activos, y con un paquete listo para ser transmitido, tengan la misma probabilidad de transmisión.

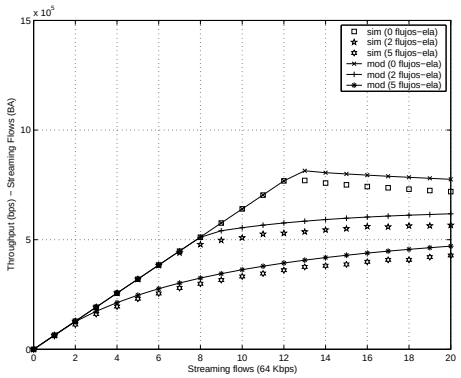


Figura 1: Caudal tráfico rígido (BA)

En la Figura 2 presentamos el caudal de los flujos elásticos en función del número de flujos rígidos. Cuando en el sistema no hay flujos rígidos, para el caso de $n_e = 2$ flujos elásticos, el caudal es superior, aunque no significativamente, al caso en que existen $n_e = 5$ flujos elásticos. Se observa como el caudal se reparte de manera equitativa entre el conjunto de nodos con flujos elásticos. Cuando no hay flujos rígidos en el sistema, la reducción del caudal con el aumento de flujos elásticos es debida a los mecanismos de acceso múltiple (aumento de colisiones). A medida que se aumenta el número de flujos rígidos, el caudal de los flujos elásticos se reduce progresivamente, observándose un punto de inflexión cuando la red entra en saturación (el coeficiente de utilización de la red es igual a $\rho_{red} = 1$). También se observa un efecto de clase donde a mayor número de flujos de un mismo tipo, esa clase de tráfico recibe un mayor caudal. Este hecho se justifica por si mismo ya que la probabilidad de acceder al medio (y por tanto transmitir), des del punto de vista de una clase de tráfico, aumenta con el número de flujos de esa clase.

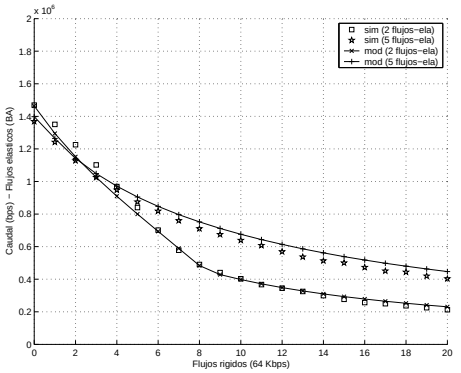


Figura 2: Caudal tráfico elástico (BA)

En la Figura 3 se muestra el caudal de los flujos elásticos cuando se fija el número de flujos rígidos, aumentando progresivamente el número de flujos elásticos. En el caso donde no hay flujos rígidos $n_s = 0$, a medida que aumentamos el número de flujos elásticos el caudal se reduce a causa del propio mecanismo de acceso aleatorio. Para los casos con $n_s = 2$ y $n_s = 5$ flujos rígidos se observa como para un único flujo elástico el caudal que obtiene es proporcional al número de flujos rígidos. A medida que aumenta el número de flujos elásticos, estos obtienen parte del

caudal de los flujos rígidos aumentando el caudal del tráfico elástico (el efecto de clase mencionado anteriormente). Esto conlleva una reducción del caudal de los flujos rígidos y por tanto, a la posibilidad que no puedan utilizar el ancho de banda que necesitan para su correcto funcionamiento.

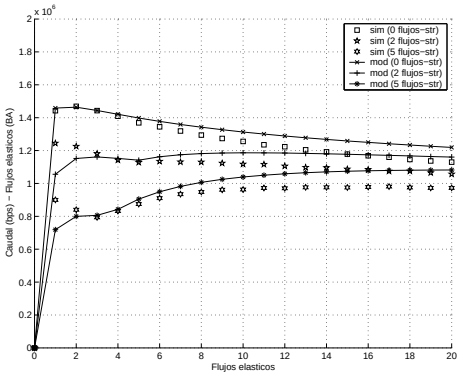


Figura 3: Caudal tráfico elástico (BA)

El parámetro clave del modelo es la utilización de la cola ($\rho^k(\mathbf{n})$) ya que captura el impacto recíproco de las propias características del tráfico con el estado de la red. En la Figura 4 se presenta la utilización de la cola para los nodos con flujos rígidos en función del número de flujos elásticos. Como se puede observar, a mayor número de flujos elásticos, la utilización de la cola en los nodos con flujos rígidos aumenta considerablemente (esto justifica los resultados presentados anteriormente). El aumento de la utilización de la cola está motivado por el aumento del tiempo de servicio, que a su vez aumenta debido al incremento de colisiones en el canal. Por tanto, la probabilidad de pérdida de paquetes en el nodo aumenta, reduciéndose el número de paquetes entregados a la red. A medida que aumentamos el número de flujos rígidos, la utilización de la cola lógicamente aumenta. Llama la atención que mientras que la utilización media de la red es reducida, la utilización de la cola aumenta muy lentamente, pasando a aumentar rápidamente cuando la red se encuentra cerca del punto de saturación. Como hemos indicado, el comportamiento de la utilización de la cola está altamente ligado a la probabilidad de colisión observada por un nodo.

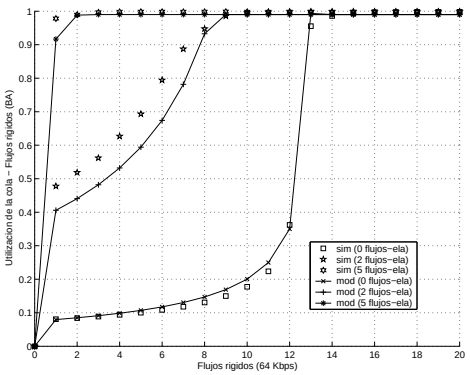


Figura 4: Utilización de la cola (BA)

3.5. Aplicaciones

La finalidad de este modelo no es la de obtener una elevada caracterización de las prestaciones del nivel MAC IEEE 802.11, sino proveer de un modelo simple que permita ser utilizado para obtener resultados rápidos de situaciones concretas o evaluar mecanismos de calidad de servicio en redes IEEE 802.11. Destacamos un conjunto de aplicaciones directas del modelo.

- Análisis de configuraciones reales de la red [12]. Utilizando el modelo se pueden definir escenarios reales ya sea en modo ad-hoc o en modo infraestructura de redes WLAN. En la siguiente sección se analizan las prestaciones de una red formada por un punto de acceso con tráfico TCP (FTP) y tráfico en tiempo real (UDP) bidireccional de 64 Kbps.
- Prestaciones de la red en entornos con diferentes velocidades de transmisión [7]. Considerando diferentes valores de R_{data} en función del estado del canal y posición del terminal móvil podemos evaluar el impacto de las diferentes velocidades de transmisión en las prestaciones de una celda WLAN. En esta situación para cada velocidad de transmisión se tendría una clase de tráfico diferente.
- Mecanismos de diferenciación de tráfico [10], [11]. El modelo permite evaluar diferentes mecanismos de diferenciación de tráfico, por ejemplo modificaciones adaptativas del valor del número de ranuras del algoritmo de espera aleatoria.
- Control de admisión. El modelo se puede utilizar para evaluar diferentes mecanismos de control de admisión en redes WLAN con tráfico heterogéneo.

4. Prestaciones de una red WLAN en modo infraestructura

En este artículo se ha aplicado el modelo para evaluar las prestaciones de una red WLAN en modo infraestructura, formada por un punto de acceso y n terminales móviles asociados. De los n terminales móviles, n_s mantienen activa una conversación bidireccional, donde cada flujo ocupa B_s bps en cada sentido de la conversación. Los n_e terminales restantes realizan una descarga de ficheros (FTP) de un servidor conectado al punto de acceso mediante un enlace de 100 $Mbps$ y 2 ms de tiempo de propagación. Para modelar el escenario descrito se utilizan tres clases de tráfico ($K = 3$), donde

1. (k=1) Tráfico descendente (desde el punto de acceso a los nodos) y formado por tráfico UDP y TCP (FTP).

2. (k=2) Tráfico ascendente UDP.

3. (k=3) Tráfico ascendente TCP (ack's).

La tasa de llegada de paquetes al punto de acceso en sentido descendente está compuesta por tráfico rígido (flujos desde la red hasta los terminales móviles) y por tráfico elástico (FTP).

$$\lambda^{k=1} = n_s\lambda_s + n_e\lambda_e \tag{25}$$

donde λ_e es la tasa de paquetes enviados por el servidor FTP hacia un terminal móvil. Para obtener $n_e\lambda_e$, se realiza la hipótesis de idealidad de los flujos elásticos (se adaptan perfectamente al estado de la red), calculándose para obtener una utilización de la cola igual a la unidad. Así, el factor de utilización del tráfico UDP es $\rho_{UDP} = n_s\frac{B_s}{L_s}X^{k=1}(\mathbf{n})$ y la tasa de paquetes de tráfico elástico

$$\lambda_e = \frac{1-(1-P_b^{k=1})\rho_{UDP}}{X^{k=1}(\mathbf{n})} \tag{26}$$

donde $P_b^{k=1}$ es la probabilidad de pérdida de paquetes en la cola del AP. Para calcular esta probabilidad y contemplar la presencia variable de flujos elásticos y rígidos se realiza la siguiente estimación del tráfico ofrecido en el enlace descendente

$$\nu^{k=1}(\mathbf{n}) = n_s\frac{B_s}{L_s}X^{k=1}(\mathbf{n}) + \frac{n_e}{n_s\rho_{UDP}+n_e} \tag{27}$$

De esta manera, en el punto de acceso (enlace descendente) se multiplexan en la misma cola $M/M/1/Q$, $n_e\lambda_e$ paquetes de longitud L_e y $n_s\lambda_s$ paquetes de longitud L_s. Para poder analizar el comportamiento de la cola, se obtiene también la longitud media de los paquetes que llegan al punto de acceso

$$E\{L^{k=1}\} = \frac{n_e\lambda_e L_e + n_s\lambda_s L_s}{n_e\lambda_e + n_s\lambda_s} \tag{28}$$

El tráfico TCP (FTP) es recibido por los n_e terminales móviles que responden a cada paquete TCP recibido (asumiendo que se reparten equitativamente) con un ACK de nivel de transporte, por tanto la tasa de transmisión de ACK's por cada nodo es $\lambda^{k=3} = \lambda_e/n_e$. Un modelado preciso del protocolo TCP tendría que considerar las fluctuaciones del tamaño de la ventana en función de los valores de RTT y de la probabilidad de pérdidas aunque, utilizando las hipótesis anteriores, se obtiene una muy buena aproximación. Si bien, este objetivo queda fuera de las pretensiones de este artículo, hace falta remarcar la importancia de un modelado correcto de la generación de tráfico para la obtención de resultados realmente precisos.

Finalmente, los terminales con conversaciones rígidas generan tráfico hacia el punto de acceso con una tasa constante de $\lambda^{k=2} = \lambda_s = B_s/L_s$ paquetes por segundo.

Los parámetros considerados son los mismos que los utilizados en apartados anteriores, cambiando la longitud de las tramas a nivel de aplicación a $L_s =$ 500 *bytes* y $L_e = 1500$ *bytes* (la longitud a nivel MAC incluye las cabeceras TCP (o UDP) y IP). También se ha modificado la tasa básica de transmisión a $R_{basic} = 1$ *Mbps* y se ha considerado la utilización del mecanismo RTS/CTS. Considerando únicamente la presencia de tráfico rígido en la red, en la Figura 5 se muestra el caudal obtenido tanto para el enlace descendente como para el enlace ascendente. Destacar la proximidad entre los resultados obtenidos utilizando el modelo y los obtenidos por simulación, en este caso a través de un escenario definido utilizando *ns-2* [14]. La diferencia entre los resultados de simulación y el modelo pueden ser causados por la utilización de tráfico CBR (no Poisson) en las simulaciones. Como se puede observar, el número de flujos de tráfico rígido (y por tanto la capacidad de la red) están limitados por el enlace descendente, en esta situación, a un valor de $n_s = 7$ flujos rígidos bidireccionales. Esto es debido a que el tráfico rígido que ha de cursar el punto de acceso en el enlace descendente es n_s veces superior al tráfico que ha de transmitir un terminal móvil pero, la probabilidad de transmisión del punto de acceso es inversamente proporcional al número de terminales móviles $(1/(n_s+1))$ ya que tanto el punto de acceso como los terminales móviles tienen la misma probabilidad de acceder al canal.

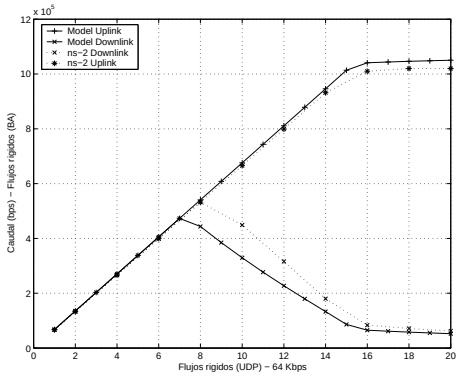


Figura 5: Caudal tráfico rígido (UDP)

En la Figura 6 se muestra el caudal en el enlace descendente en presencia de tráfico TCP y UDP en función del número de flujos rígidos activos. Como se puede observar, el caudal de tráfico elástico se reduce a medida que aumenta el número de flujos de tráfico rígido. Este comportamiento está motivado por que tanto el tráfico UDP como TCP comparten una misma cola y el tráfico elástico reacciona ante la pérdida de paquetes y al aumento del RTT. Así mismo, se observa como a mayor número de flujos elásticos la reducción del caudal de tráfico TCP es menor. Finalmente, en la Figura 7 podemos observar el caudal de tráfico UDP en el enlace descendente. Tanto los resultados de simulación como los obtenidos utilizando el modelo muestran que a mayor número de flujos elásticos, como era de esperar, el caudal del tráfico UDP se reduce y por tanto la capacidad de la red (en términos de flujos rígidos) queda afectada por el número de flujos elásticos.

Uno de los aspectos más complejos, y en el que se continuará trabajando, es el análisis de como los

diferentes flujos de tráfico TCP y UDP comparten la cola situada en el punto de acceso. En este artículo se ha considerado un comportamiento ideal del tráfico TCP, calculando la tasa de tráfico elástico en función de la utilización de la cola que requerían los flujos de tráfico rígido. Esta aproximación ha de ser refinada para poder contemplar con mayor precisión la presencia de múltiples flujos TCP (no ideales) y la consiguiente degradación de las prestaciones para el tráfico UDP.

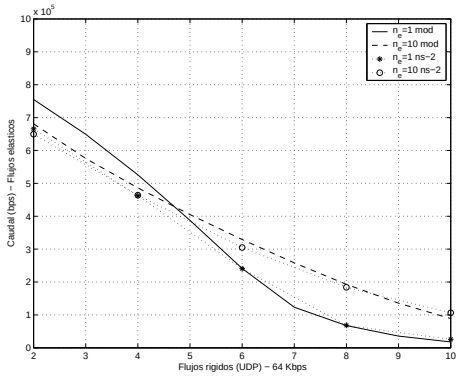


Figura 6: Caudal tráfico elástico (TCP)

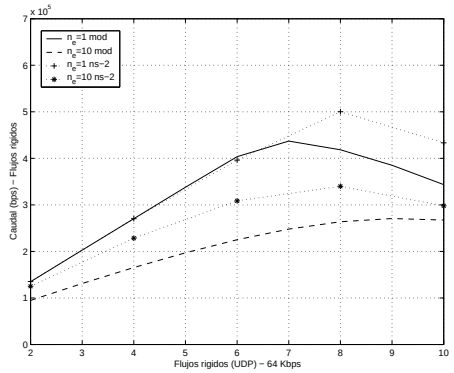


Figura 7: Caudal tráfico rígido (UDP)

5. Conclusiones

En esta comunicación se ha presentado un modelo analítico del protocolo MAC IEEE 802.11 que permite evaluar la interacción de flujos de tráfico con diferentes características en una red WLAN. El modelo puede ser aplicado de manera directa en multitud de escenarios, especialmente los relacionados con el diseño de mecanismos de calidad de servicio.

De los resultados presentados destacar los efectos que tiene la inclusión de tráfico elástico en las prestaciones del tráfico rígido (a causa del acceso múltiple), normalmente asociado a aplicaciones de tiempo real como aplicaciones de transmisión de voz, vídeo o ambas. Para proteger estos flujos de tráfico del efecto de los flujos de tráfico elástico será necesario introducir mecanismos de diferenciación de tráfico y de control de admisión.

Referencias

[1] IEEE 802.11 WG, Part 11; *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications*, ANSI/IEEE Std 802.11-1999.

[2] G. Bianchi; *Performance Analysis of the IEEE 802.11 Distributed Coordination Function*, IEEE Journal on Selected Areas in Communications, Vol. 18, No. 3, March 2000.

[3] Marcelo M. Carvalho, J.J. Garcia-Luna-Aceves; *Delay Analysis of IEEE802.11 in Single-Hop Networks*, 11th IEEE International Conference on Network Protocols (ICNP'03), Atlanta, Georgia, November 2003.

[4] Omesh Tickoo and Biplab Sikdar; *Queuing Analysis and Delay Mitigation in IEEE 802.11 Random Access MAC based Wireless Networks*, IEEE INFOCOM, Hong Kong, China, March 2004.

[5] Gion Reto Cantieni, Qiang Ni, Chadi Barakat, Thierry Turletti; *Performance Analysis under Finite Load and Improvements for Multirate 802.11*, Preprint submitted to Elsevier Science, June 2004.

[6] David Malone, Ken Duffy, Douglas J. Leith; *Modeling the 802.11 distributed Coordination Function with Heterogenous Finite Load*, Resource Allocation in Wireless Networks, April 3rd, 2005, Trento, Italy.

[7] Martin Heusse, Franck Rousseau, Gilles Berger-Sabbatel, Andrzej Duda; *Performance Anomaly of 802.11b*, Proceedings of the IEEE Infocom 2003.

[8] F. Delcoigne, A Proutière, G. Régnié; *Modelling integration of streaming and data traffic*, Elsevier Science Publishers B. V. 2004.

[9] Saar Pilosof, Ramachandran Ramjee, Danny Raz, Yuval Shavitt, and Prasun Sinha; *Understanding TCP fairness over Wireless LAN*, Proceedings of IEEE Infocom , Apr 2003.

[10] IEEE 802.11 WG; *Part 11: Wireless Medium Access Control (MAC) and physical layer (PHY) specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS)*, em IEEE 802.11e/D2.0,Nov. 2001.

[11] A. Banchs, X. Pérez, M. Radimirsch, H.J.Stüttgen; *Service Differentiation Extensions for Elastic and Real-Time traffic in 802.11 Wireless LAN*, In Proceedings of the IEEE Conference on High Performance Switching and Routing (HPSR 2001), Dallas, USA, May 2001

[12] R. Bruno, M. Conti, E. Gregori; *Analytical Modeling of TCP Clients in Wi-Fi Hot Spot Networks*, Networking 2004. LNCS 3042, pp.626-637, 2004.

[13] Gilbert (Gang) Chen; *Component Oriented Simulation Toolkit*, http://www.cs.rpi.edu/ cheng3/

[14] *Ns2 Network Simulator*, http://www.isi.edu/nsnam/ns/

Estudio del impacto de los parámetros de configuración de los protocolos AODV y OLSR en entornos reales MANET

C. Gómez, J. Paradells
Wireless Networks Group, Departamento de Ingeniería Telemática, Universitat Politècnica de Catalunya
Jordi Girona 1-3, C-3, 08034 Barcelona, España
Teléfono: 93 413 72 06 Fax: 93 413 70 07
E-mail: {carlesgo, teljpa}@entel.upc.edu

Abstract. *Mobile Ad-hoc NETworks (MANETs) have attracted the attention of a large number of researchers in the last years. While routing has been the main research activity in this field, surprisingly, little effort has been devoted up to now to analyzing the impact of routing protocol parameters on performance of ad-hoc networks. In this paper we study the influence of the routing protocol, its implementation and its configuration on network performance using AODV and OLSR. Results suggest that usage of different settings from those proposed by default increase reactivity after topology changes while having a small impact on bandwidth and power consumption in a set of scenarios.*

1 Introducción

Las redes móviles ad-hoc (MANETs) han atraído la atención de una gran cantidad de investigadores en los últimos diez años debido a avances significativos de las tecnologías inalámbricas en cuanto a estandarización y disponibilidad de las mismas, junto con la popularidad del uso de dispositivos de tamaño pequeño (desde portátiles hasta sensores de un milímetro cúbico). Las tecnologías MANET permiten la comunicación a un conjunto de nodos inalámbricos sin requerir la existencia de una infraestructura de red previa. Esta característica posibilita la creación de redes aisladas (por ejemplo, para aplicaciones de rescate y emergencia) y es también una pieza esencial en el camino hacia una Internet ubicua, por ejemplo, permitiendo extender la cobertura de redes inalámbricas con infraestructura.

Sin embargo, las redes ad-hoc sufren limitaciones en el rendimiento que las comunicaciones de datos pueden lograr en estos entornos, debido a la movilidad de los nodos, problemas inherentes a la transmisión radio, autonomía reducida debido a la batería de los nodos y la falta de infraestructura de red. Un elemento clave con influencia en la eficiencia de la red es el protocolo de encaminamiento. Idealmente, un protocolo de encaminamiento de red ad-hoc debería ser capaz de proporcionar rutas óptimas en el menor tiempo posible, incluso en el caso de roturas de enlaces en un camino activo (es decir, con una comunicación en curso), con un mínimo impacto en la latencia entre origen y destino, el ancho de banda disponible y el consumo de batería de los dispositivos para cualquier patrón de tráfico. Dado que la operación de los protocolos de encaminamiento puede ser regulada a través de los valores de sus parámetros, el uso de una configuración de protocolos adecuada es un factor importante para optimizar el rendimiento de la red.

Sorprendentemente, hasta ahora existe poco trabajo dedicado al análisis del impacto de los parámetros de los protocolos de encaminamiento en redes ad-hoc en el rendimiento de las mismas. Por otra parte, las métricas de rendimiento usadas en la literatura suelen ser el porcentaje de paquetes entregados y la latencia extremo a extremo, mientras que no se ha considerado hasta el momento el análisis de un parámetro crítico como la latencia de cambio de ruta (es decir, tiempo que tarda un protocolo de encaminamiento en hallar una ruta alternativa tras la rotura de un enlace activo) y su dependencia con los parámetros de los protocolos de encaminamiento.

En este artículo presentamos nuestra experiencia testeando el efecto de la configuración de parámetros de protocolos de encaminamiento en escenarios reales de red ad-hoc utilizando el Ad-hoc On-demand Distance Vector (AODV) [1] y el Optimized Link State Routing (OLSR) [2], actualmente los dos protocolos de encaminamiento MANET con el mayor número de implementaciones disponibles de forma pública. Nuestro objetivo no es comparar ambos protocolos, hecho que requeriría un análisis en una amplio espectro de situaciones distintas, estudiando el efecto de factores como el patrón de tráfico, número de nodos y fuentes, topología de red, patrón de movilidad o tipo de dispositivos, por citar algunos de los más relevantes. Nuestra intención es analizar de forma teórica y empírica el impacto de los parámetros de configuración de estos protocolos sobre un conjunto de métricas de rendimiento en maquetas con dispositivos reales. Tales métricas son: retardo extremo a extremo, ancho de banda extremo a extremo, latencia de cambio de ruta y consumo de batería. Asimismo, los experimentos realizados permitirán cuantificar y/o indicar efectos derivados del uso de implementaciones reales de AODV y OLSR sobre los parámetros de rendimiento indicados. Cabe añadir que tampoco existen en la

literatura resultados acerca del consumo de batería en dispositivos reales a causa de tales protocolos.

El resto del artículo se estructura de la siguiente forma. La sección 2 se dedica a una visión general del protocolo AODV, donde se focaliza en los principales mecanismos y parámetros que pueden influir en el rendimiento de una red ad-hoc. Similarmente, presentamos y discutimos la operación de OLSR en la sección 3. Describimos los escenarios donde se realizan nuestros experimentos en la sección 4, explicando nuestra metodología y los resultados logrados en la sección 5. Finalmente, en la sección 6 resumimos las conclusiones más relevantes de este trabajo, indicando direcciones futuras de investigación que de él se derivan.

2 AODV: mecanismos y parámetros

2.1 Visión general

De acuerdo con la naturaleza reactiva de AODV, cuando un nodo requiere una ruta (nos referiremos a este nodo como "originador"), inicia un procedimiento de descubrimiento de ruta mediante la difusión de mensajes Route Request (RREQ). Cuando un nodo recibe un RREQ, si dispone de una entrada válida en su tabla de encaminamiento para el destino requerido, o es él mismo el destino, entonces genera y manda un mensaje Route Reply (RREP) hacia el nodo que ha originado el descubrimiento de ruta. Cada nodo mantiene entradas en la tabla de encaminamiento con información sobre el siguiente salto que expira después de un tiempo si el camino se convierte en inactivo.

Cuando un enlace de un camino activo se rompe, el nodo más cercano al originador que detecta la rotura del enlace crea un mensaje Route Error (RERR) que reporta el conjunto de destinos que a partir de este momento ya no serán alcanzables y lo manda a sus nodos precursores. El originador puede entonces iniciar un nuevo descubrimiento de ruta, o bien, otra opción es que los nodos intermedios pueden reparar de forma local la ruta.

2.2 Mecanismos de detección de rotura de enlaces

Un factor clave en el rendimiento de un protocolo de encaminamiento consiste en cómo este protocolo mantiene información acerca de la existencia o disponibilidad de enlaces en la red. Un nodo AODV que forma parte de un camino activo puede difundir mensajes Hello de forma local (es decir, a sus vecinos en un radio de un salto). AODV define también otras estrategias de detección de rotura de enlace, como la notificación a nivel de enlace (por ejemplo, la ausencia de un reconocimiento de nivel 2) el denominado reconocimiento pasivo (escuchar el reenvío de un paquete por parte del siguiente salto de un nodo) o bien la recepción de cualquier otro paquete por parte del siguiente salto de un nodo, incluyendo respuestas a RREQ o ICMP Echo Request.

La mayoría de implementaciones de AODV usan por defecto el mecanismo de detección de caída de enlace basado en el envío de mensajes Hello. Una razón es que la notificación por parte del nivel de enlace, evidentemente, depende de la tecnología de nivel 2 utilizada. Por tanto, el uso de mensajes Hello no restringe una implementación a una tecnología de nivel 2 específica. Por otra parte, de esta forma se evita la complejidad que conlleva la interacción entre capas.

2.3 Parámetros relevantes

En este trabajo centramos nuestra atención en los parámetros de AODV que regulan el mecanismo de detección de rotura de enlace mediante mensajes Hello: ALLOWED_HELLO_LOSS y HELLO_INTERVAL. El tiempo máximo entre la transmisión de dos mensajes Hello es igual a HELLO_INTERVAL ms. Después de recibir un mensaje Hello por parte de un vecino, si un nodo no recibe ningún paquete del mismo durante más de ALLOWED_HELLO_LOSS*HELLO_INTERVAL milisegundos, el nodo debe asumir que el enlace está roto. Por defecto, ALLOWED_HELLO_LOSS es igual a 2 y HELLO_INTERVAL es igual a 1000 ms [1]. En la sección 5 se mostrará el efecto de un rango de valores de HELLO_INTERVAL en el rendimiento de nuestros escenarios ad-hoc.

2.4 Efecto de AODV en el rendimiento de la red ad-hoc

A continuación se discute el efecto de AODV y los parámetros del protocolo que consideramos en el retardo extremo a extremo, el ancho de banda disponible, la latencia de cambio de ruta y el consumo de batería de los nodos de una red ad-hoc.

1) Retardo extremo a extremo. Como protocolo reactivo, AODV introduce una Latencia de Descubrimiento de Ruta (LDR) en el primer paquete de una transmisión de datos si no existe una entrada válida para el destino. La LDR dependerá del número de saltos extremo a extremo. Por otra parte, también cabe esperar que el procesado de los mensajes de RREQ y RREP en los nodos intermedios contribuya a la LDR.

2) Ancho de banda disponible extremo a extremo. Nos referimos al ancho de banda extremo a extremo alcanzable en un camino que no sufre ninguna rotura de enlace, un evento que consideramos en el parámetro de rendimiento presentado a continuación. El parámetro HELLO_INTERVAL tiene un impacto directo en el ancho de banda disponible de un camino, que decrece a medida que crece la frecuencia de envío de mensajes de control. En nuestros experimentos cuantificaremos este efecto.

3) Latencia de Cambio de Ruta (LCR). Definimos este parámetro como el tiempo total entre el instante en el cual se rompe un enlace en un camino activo y el momento en que el originador empieza a usar una ruta alternativa, si ésta existe. La LCR depende del producto ALLOWED_HELLO_LOSS*HELLO_INTERVAL. Evidentemente, reducir ALLOWED_HELLO_LOSS manteniendo una frecuencia de mensajes Hello fija ayudaría a decrementar la LCR. Sin embargo, dado que el valor por defecto de ALLOWED_HELLO_LOSS es igual a 2, la única posibilidad consistiría en configurar este parámetro con un valor igual a 1. En este caso, el rendimiento podría decrecer si malas condiciones radio resultan en pérdidas aisladas de mensajes Hello, hecho que llevaría a detecciones espúreas de roturas de enlaces. Por otra parte, se ha demostrado mediante experimentos de campo que configurar ALLOWED_HELLO_LOSS con un valor igual a 3 decrementa el rendimiento, debido a que la reactividad frente a cambios en la topología se reduce [3]. Por tanto, el valor usado por defecto es una buena elección. Por su parte, consideramos que HELLO_INTERVAL es un buen candidato para regular la LCR. Para minimizar este parámetro, es deseable que se usen valores de HELLO_INTERVAL bajos, hecho que pone de manifiesto un compromiso con la disponibilidad de ancho de banda.

A continuación, caracterizamos analíticamente la LCR. Consideremos en primer lugar un caso simple donde el enlace entre un nodo activo, que denominaremos A, y su siguiente salto falla (por ejemplo, el siguiente salto se apaga, se desplaza fuera de la zona de cobertura del primer nodo, etc.). Asumimos un valor de ALLOWED_HELLO_LOSS igual a 2. Por tanto, después de un periodo igual a 2*HELLO_INTERVAL durante el cual no se reciben mensajes Hello, el nodo A decide que el enlace se ha roto. Podemos plantear dos situaciones extremas (ver Fig. 1): a) la rotura del enlace ocurre inmediatamente después de la recepción del último mensaje Hello del anterior vecino; b) el enlace se rompe casi un HELLO_INTERVAL después de la recepción del último mensaje Hello. Sea T_{Detec} el tiempo entre el momento en que ocurre la rotura de enlace y el instante en que ésta es detectada por el nodo A. Dado que el instante de rotura del enlace es a priori desconocido, podemos caracterizar T_{Detec} como una variable aleatoria uniformemente distribuida entre HELLO_INTERVAL y 2*HELLO_INTERVAL. La LCR puede ser calculada teniendo en cuenta que los valores de LDR son pequeños comparados con T_{Detec} (ver sección 5) como:

$$LCR = T_{Detec} + LDR \approx T_{Detec} \quad (1)$$

4) Consumo de batería. Incrementar la frencuencia de envío de mensajes Hello reduce el tiempo de vida de la batería de un dispositivo. Cuantificaremos este efecto en la sección 5. Debe tenerse en cuenta que en algunos casos, un nodo puede ser estático y conectado a la red eléctrica, como por ejemplo, un ordenador de sobremesa.

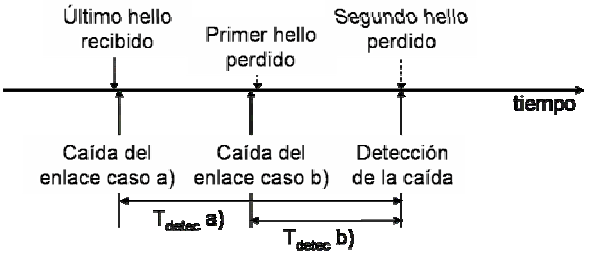


Fig. 1. Tiempo de detección de la caída de un enlace

Hemos señalado que reducir el valor de HELLO_INTERVAL provoca un mayor consumo de ancho de banda y batería, mientras que la LCR decrece. Existe, por tanto, un compromiso entre estas métricas de rendimiento. La realización de experimentos en nuestros escenarios reales permitirá obtener conclusiones acerca de valores adecuados para el HELLO_INTERVAL en un conjunto de escenarios.

3 OLSR: parámetros y mecanismos

3.1 Visión general

El protocolo OLSR es una optimización del concepto de encaminamiento de estado del enlace clásico, que se fundamenta en el uso de nodos especiales denominados MultiPoint Relays (MPRs). Los MPRs de un nodo A garantizan que todos los vecinos del nodo A en un radio de dos saltos puedan recibir los mensajes de estado del enlace, que deben ser difundidos a toda la red. Únicamente los MPRs pueden realizar esta tarea de difusión. Por tanto, el tráfico de encaminamiento se ve reducido significativamente en comparación con el método de inundación básico. Además, los MPRs pueden anunciar únicamente los enlaces entre ellos mismos y sus selectores (nodos que les han elegido como MPRs), resultando en un mecanismo de estado del enlace parcial. Estos anuncios de estado del enlace, generados periódicamente por los MPRs, son conocidos como los mensajes Topology Control (TC). Por otra parte, los nodos realizan tareas de sensado del enlace mediante la difusión periódica de mensajes Hello en un radio de un salto. Estos mensajes incluyen una lista de los vecinos con los cuales existe un enlace, proporcionando la información necesaria para poder llevar a cabo la elección de MPRs.

3.2 Parámetros relevantes

La especificación de OLSR [2] define un conjunto de constantes que regulan el uso de los mecanismos de OLSR. Esta especificación determina que la configuración de parámetros del protocolo se puede

realizar de forma independiente en cada nodo y propone un conjunto de valores por defecto. A continuación, presentamos un resumen de los parámetros más relevantes que pueden afectar al rendimiento de la red.

1) HELLO_INTERVAL. Este parámetro determina el tiempo entre mensajes Hello enviados. Se incluye en los mensajes Hello, de modo que un nodo conoce el instante en el cual debe producirse la llegada del siguiente Hello procedente del mismo nodo.

2) *REFRESH_INTERVAL*. Los mensajes Hello pueden ser parciales para minimizar el tráfico de encaminamiento. Cada enlace y cada vecino deben ser anunciados al menos una vez en un periodo REFRESH_INTERVAL. Por tanto, el parámetro HELLO_INTERVAL debe ser menor o igual al REFRESH_INTERVAL.

3) TC_INTERVAL. Los mensajes TC pueden ser parciales como los mensajes Hello. TC_INTERVAL es el periodo en el cual el conjunto de enlaces anunciados debe ser completo.

4) NEIGHB_HOLD_TIME. Indica el tiempo durante el cual la información proporcionada por un mensaje Hello debería ser considerada válida.

5) TOP_HOLD_TIME. Este parámetro es análogo al anterior y define el periodo de validez para la información de un mensaje TC.

6) WILLINGNESS. Define la voluntad de un nodo de actuar como encaminador de tráfico. Se especifica entre un conjunto de ocho posibles valores.

La Tabla 1 muestra los valores por defecto propuestos en [2] para las constantes indicadas. Otros parámetros que pueden tener influencia en el rendimiento de OLSR son los siguientes: MPR_COVERAGE, que permite definir la cantidad de MPRs que deberían cubrir a un nodo; TC_REDUNANCY, que ajusta el conjunto de enlaces anunciados en un mensaje TC y, finalmente, parámetros que regulan un mecanismo de histéresis de enlace, que afectan al establecimiento y caída de enlaces.

TABLA 1
VALORES POR DEFECTO DE LOS PRINCIPALES PARÁMETROS DE OLSR

Parámetro	Valor por defecto
HELLO_INTERVAL	2 s
REFRESH_INTERVAL	2 s
TC_INTERVAL	5 s
NEIGHB_HOLD_TIME	3 * REFRESH_INTERVAL
TOP_HOLD_TIME	3 * TC_INTERVAL
WILLINGNESS	3

3.3 Discusión

A continuación discutimos el efecto de los parámetros mencionados de OLSR sobre cada una de las métricas de rendimiento de una red ad-hoc ya consideras en el punto 3.4.

1) Retardo extremo a extremo. No existe relación directa entre ningún parámetro de OLSR y el retardo extremo a extremo de un camino dado. Sin embargo, el parámetro WILLINGNESS podría producir un efecto indirecto sobre este parámetro de rendimiento, dado que los nodos con valores de WILLINGNESS más elevados son susceptibles de sufrir congestión y, por tanto, incrementar los tiempos de espera en cola de los paquetes.

2) Ancho de banda extremo a extremo. Los parámetros HELLO_INTERVAL (que asumimos adecuadamente ajustado al valor REFRESH_INTERVAL), y TC_INTERVAL tienen un impacto directo en este parámetro de rendimiento. La capacidad disponible de un camino decrece con la frecuencia de envío de mensajes de control.

3) Latencia de Cambio de Ruta (LCR). Para minimizar la LCR, HELLO_INTERVAL y TC_INTERVAL deberían, idealmente, tener valores pequeños. En consecuencia, existe un compromiso con el ancho de banda extremo a extremo de un camino. Por otro lado, la LCR crece con parámetros como NEIGHB_HOLD_TIME y TOP_HOLD_TIME. Sin embargo, estos parámetros deberían ser mayores que el HELLO_INTERVAL y TC_INTERVAL para evitar cambios de ruta espúreos debidos a la pérdida de mensajes de control a causa de condiciones pobres de propagación radio. De este modo, existe también un compromiso para los valores de parámetros de *hold time*. Con razonamientos similares a los empleados para determinar el valor de la LCR con AODV, pero aplicados a los mensajes TC y Hello, podemos deducir que, dadas las frecuencias de envío de mensajes de control, y en particular, la de los mensajes TC y sus correspondientes *hold times*, la LCR con OLSR resultará superior en, aproximadamente, un orden de magnitud a la que se tiene con AODV.

4) Consumo de batería. Varios parámetros de OLSR afectan al tiempo de vida de la batería de un nodo. El consumo de batería crece con la frecuencia de envío de mensajes de control, por tanto, HELLO_INTERVAL y TC_INTERVAL tienen también un impacto en la supervivencia de la red. Sin embargo, aparece otro compromiso, dado que tales parámetros deberían ser pequeños para minimizar la LCR. Finalmente, WILLINGNESS es otro parámetro relevante, puesto que un valor elevado para este parámetro incrementa la probabilidad de que un nodo sea elegido como MPR y de que deba realizar tareas de encaminamiento de tráfico.

Hemos identificado distintos compromisos entre parámetros de OLSR. Nuestros tests en escenarios reales permitirán cuantificar los efectos de un conjunto de configuraciones de parámetros de OLSR distintas, además de proporcionar algunas

indicaciones con el objetivo de optimizar globalmente el rendimiento de una red ad-hoc.

4 Escenarios de pruebas

Esta sección describe el equipamiento hardware y software empleado en nuestros experimentos, así como las topologías de red utilizadas. Es importante notar que pretendemos extraer conclusiones y cuantificar resultados a partir del uso de un conjunto de configuraciones de parámetros de protocolos sobre esquemas reales que pueden constituir un marco de referencia para el análisis de esquemas más complejos.

4.1 Equipamiento

1) Dispositivos y tarjetas de red. Las pruebas con el protocolo OLSR han sido realizadas con redes formadas por portátiles Acer 662 LCi con Linux Fedora Core 2 (que actúan como extremos de la comunicación) y PDAs Zaurus SL-550S (que actúan como nodos intermedios). Las pruebas con AODV sólo han podido ser realizadas sobre portátiles debido a la falta de soporte para plataforma PDA en el momento de realización de los experimentos. En 4.2 se describen las topologías empleadas. Se han empleado tarjetas IEEE 802.11b con el mecanismo RTS/CTS deshabilitado.

2) Implementaciones de AODV y OLSR. Empleamos la implementación de AODV de Uppsala University, v.0.81 [4] y la implementación de OLSR de Unik v.0.4.0 [5]. Ambas han sido utilizadas en trabajos publicados basados en experimentos reales [3, 6]. La primera permite la configuración de todos los parámetros de AODV mediante recompilación de su código, mientras que la segunda permite que el usuario defina tales valores sin recompilar código. La mayor parte de ambas implementaciones reside en un demonio de espacio de usuario, por lo cual, dado que los paquetes deben atravesar el kernel hasta el espacio de usuario, se espera que su rendimiento se vea afectado si se compara con otras posibilidades de implementación, como por ejemplo, el uso de código en el kernel que permita disparar los eventos de AODV u OLSR [9].

3) Herramientas de generación de tráfico. Empleamos la herramienta iperf [7] para generar tráfico UDP y TCP y realizar medidas en el lado del extremo receptor, capturando el tráfico en ambos extremos mediante Ethereal [8].

4.2 Escenarios

Usamos dos tipos de esquemas para realizar nuestras medidas. A continuación, se presenta cada uno de ellos.

1) Topología en cadena de N saltos. Esta topología consiste en un conjunto de N+1 nodos alineados, donde N varía entre 1 y 4 saltos. Los nodos permanecen estáticos en este escenario. Los más distantes actúan como extremos de la comunicación.

2) Topología de N caminos de 2 saltos. Esta topología consiste en dos portátiles comunicándose a través de uno de un conjunto de N caminos de dos saltos posibles, de modo que tanto el emisor como el receptor disponen de N vecinos candidatos a actuar como encaminadores de su tráfico. En este escenario, concretamente para N=2, se analiza el efecto de la rotura de un enlace mediante el apagado del nodo activo que actúa como conmutador de paquetes.

Vale la pena comentar que controlamos la conectividad entre nodos usando la herramienta iptables para efectuar filtraje por MAC. Este enfoque es distinto al de un escenario de campo, donde el rendimiento de las comunicaciones está más sujeto a degradación por distintos motivos. Sin embargo, debe tenerse en cuenta que nuestro objetivo no requiere la presencia de ese efecto, que ha sido analizado adecuadamente en trabajos como [3, 6]. Por tanto, elegimos realizar nuestros experimentos en un escenario más controlable.

5. Experimentos y resultados

A continuación presentamos nuestra definición de experimentos junto con los resultados asociados. Nuestro objetivo es evaluar el retardo extremo a extremo, el ancho de banda disponible, la LCR y el consumo de batería en nuestros esquemas ad-hoc, analizando la influencia de la operación de los protocolos de encaminamiento y la configuración de sus parámetros en el citado conjunto de métricas de rendimiento.

5.1 Retardo extremo a extremo

Hemos realizado medidas de retardo extremo a extremo ejecutando 15 veces un experimento consistente en mandar 10 paquetes Echo Request mediante la aplicación ping en el escenario con topología en cadena de N saltos, donde N varía de 1 a 4, con AODV y con OLSR. Para las pruebas con AODV, nos aseguramos de que las entradas de las tablas de encaminamiento han expirado antes de realizar un nuevo experimento, para poder medir la LDR del primer paquete.

La Fig. 2 muestra la media de los tiempos de ida y vuelta (RTTs). En el caso de AODV, distinguimos la media de los 15 RTTs que incluyen LDR del resto de RTTs. Denotamos a los primeros como RTT_LDR/AODV y a los segundos como RTT/AODV. Indicamos los RTTs obtenidos con OLSR como RTT/OLSR.

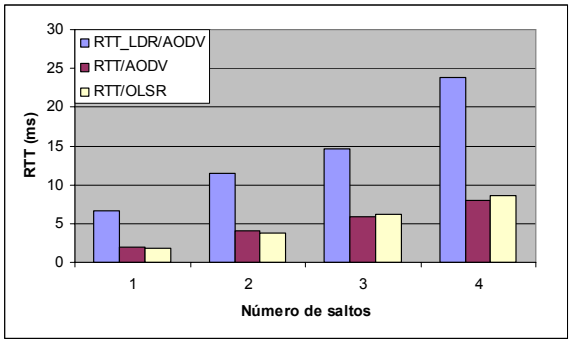


Fig. 2. Medidas de RTT en una topología en cadena de N saltos, donde N varía entre 1 y 4 saltos

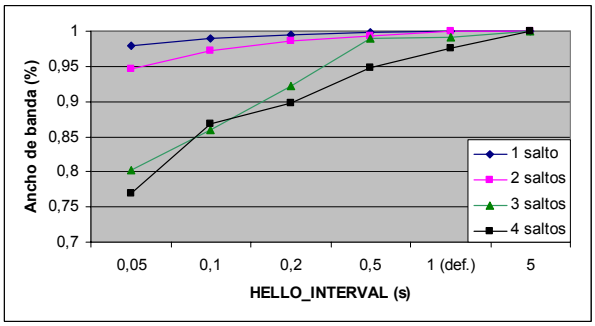


Fig. 3. Ancho de banda extremo a extremo normalizado en una topología en cadena de N saltos donde N varía entre 1 y 4 saltos, para distintos valores de HELLO_INTERVAL de AODV

Uno de los resultados más relevantes es que, pese a que los mensajes RREQ o RREP de AODV y los paquetes Echo Request o Echo Reply disponen de tamaños comparables, los RTTs que incluyen un componente de LDR son mayores que 2*RTT. La causa radica en que AODV se ejecuta como un demonio de espacio de usuario, hecho que introduce un retardo significativo en el *forwarding* o reenvío de RREQ/RREP en los nodos intermedios y en el procesado de RREQ/RREP en los extremos finales de la comunicación. De acuerdo con nuestras medidas, el retardo de reenvío de RREQ/RREP varía entre 0.7 ms y 1 ms. El procesado de un RREQ requiere cerca de 1 ms, mientras que el de un RREP tarda desde 0.6 ms hasta 5.5 ms. Dado que cada salto añade unos 2 ms al RTT, la operación de la implementación de AODV introduce retardos que no pueden despreciarse.

Como cabía esperar, OLSR no añade ningún retardo adicional al tiempo de ida y vuelta, puesto que si se conoce una ruta hasta el destino, se utiliza inmediatamente. Finalmente, cabe comentar que los RTT crecen linealmente con el numero de saltos.

5.2 Ancho de banda

A continuación presentamos un conjunto de experimentos efectuados con AODV y OLSR, con el objetivo de medir el efecto de los parámetros de ambos protocolos en el ancho de banda disponible entre los extremos de la comunicación.

En el primer caso, testeamos el impacto de la frecuencia de envío de mensajes Hello en el ancho de banda disponible extremo a extremo en una red con topología en cadena de *N* saltos, donde *N* varía entre 1 y 4. Transmitimos un flujo UDP entre los dos extremos a una tasa ligeramente superior a la capacidad disponible en cada escenario, para un conjunto de valores de HELLO_INTERVAL.

La Fig. 3 ilustra la media de 15 pruebas para cada caso. Como se puede apreciar, la presencia de mensajes Hello sólo tiene un efecto relevante (rendimiento inferior al 95%) en el ancho de banda disponible para los escenarios con topología en cadena de 3 y 4 saltos cuando los valores de HELLO_INTERVAL usados son inferiores a 0.5 segundos. Dos causas llevan a los resultados obtenidos: i) colisiones entre mensajes Hello y datos y ii) la incapacidad de enviar mensajes Hello en los instantes adecuados, especialmente para los valores de HELLO_INTERVAL más pequeños, que dan lugar a intervalos entre mensajes Hello mayores a los esperados, hecho que provoca detecciones espúreas de roturas de enlaces.

En el caso de OLSR, se obtienen resultados similares. Por ejemplo, en el caso de una topología de 2 caminos de 2 saltos, el uso de la configuración #1 da lugar a un 94% del ancho de banda disponible, mientras que se tiene un 97% del mismo cuando se usa la configuración por defecto, pero el número de caminos de 2 saltos es igual a 4 (de forma que hay 4 vecinos generando mensajes periódicos de control).

5.3 Latencia de cambio de ruta

Nuestras medidas de la LCR se llevan a cabo en el escenario de 2 caminos de 2 saltos apagando el nodo intermedio empleado en el camino activo, forzando el uso de la ruta alternativa. Debe notarse que esta acción puede emular también el desplazamiento fuera del área de cobertura del emisor por parte de su siguiente salto. Realizamos estas pruebas mandando un flujo de paquetes UDP, con un tiempo entre paquetes enviados igual a 2 ms. Nos centramos en el efecto de la LCR en el lado del receptor, midiendo el

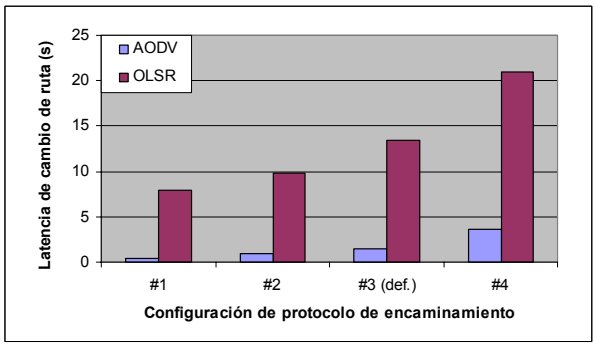


Fig. 4. Efecto de la frecuencia de envío de mensajes de control sobre la LCR

TABLA 2
CONFIGURACIONES DE OLSR EMPLEADAS EN LOS EXPERIMENTOS

Parámetros de OLSR	#1	#2	#3	#4
HELLO_INTERVAL (s)	0.5	1	2	4
REFRESH_INTERVAL (s)	0.5	1	2	4
TC_INTERVAL (s)	1.25	2.5	5	10
NEIGHB_HOLD_TIME (s)	1.5	3	6	12
TOP_HOLD_TIME (s)	3.75	7.5	15	20
WILLINGNESS	AUT.	AUT.	AUT.	AUT.

TABLA 3
CONFIGURACIONES DE AODV PARA LAS PRUEBAS DE LCR

Parámetro de AODV	#1	#2	#3	#4
HELLO_INTERVAL (s)	0.25	0.5	1	2

correspondiente agujero de conectividad como el tiempo total entre el instante en que el último paquete se recibe antes de la caída del enlace y el instante en que el primer paquete tras el cambio de ruta es recibido. La Tabla 2 muestra el conjunto de configuraciones de parámetros de OLSR considerados en estas pruebas. Para las pruebas de AODV se ha adoptado una estrategia similar, usando las configuraciones mostradas en la Tabla 3. En la Fig. 4, se muestra la media de cinco pruebas usando cada una de las configuraciones probadas.

Como cabía esperar, la LCR alcanza valores significativos con OLSR. Por ejemplo, con su configuración por defecto (configuración #3), el agujero de conectividad medio percibido en el receptor es de 13 segundos, unas diez veces superior a lo que ocurre con el uso de AODV con su configuración por defecto. Este resultado confirma que OLSR no es un protocolo adecuado para entornos con alta movilidad, dada su baja reactividad a cambios en la topología. De todos modos, tal como muestra la Fig. 4, se puede lograr una reducción de hasta 5 s incrementando en un factor igual a 4 la frecuencia de emisión de mensajes de control, con respecto al uso de los parámetros por defecto de OLSR. En el caso de AODV, los agujeros de conectividad medidos caen en el margen definido entre HELLO_INTERVAL y 2*HELLO_INTERVAL. También hemos realizado experimentos con valores de HELLO_INTERVAL inferiores a 0.25 segundos.En tales casos, la implementación de AODV no logra reaccionar adecuadamente a la ausencia de mensajes Hello tras un periodo igual a ALLOWED_HELLO_LOSS*HELLO_INTERVAL.

En el resto de casos, la LCR medida corresponde a los valores predichos por (1). Cabe notar que la LCR medida incluye un componente de LDR, pero éste es despreciable frente a T_{Detec}, de acuerdo con los valores mostrados en la Fig. 4.

5.4 Consumo de batería

Finalmente, cuantificamos los efectos sobre el consumo de batería de la configuración de AODV y OLSR.

En el primer caso, realizamos experimentos en un escenario con topología en cadena de 2 saltos. Las medidas se efectúan en el portátil intermedio, mientras una transmisión TCP se lleva a cabo entre los extremos de la comunicación. La batería del portátil objeto de las medidas está cargada completamente al inicio de cada experimento, el cual termina cuando sólo resta un 3% de batería. La pantalla se apaga automáticamente tras 60 segundos de inactividad. La Fig. 5 muestra los resultados de estas pruebas. La principal conclusión es que el consumo de batería no se ve significativamente afectado por la reducción del parámetro HELLO_INTERVAL. Por ejemplo, el envío de mensajes Hello cada 250 ms resulta en un tiempo de vida de la batería igual al 96.9% del obtenido con la configuración por defecto.

Para el caso de OLSR, empleamos el escenario de topología con 2 ramas y 2 saltos, para el conjunto de configuraciones de parámetros de OLSR mostradas en la Tabla 2, salvo para el parámetro WILLINGNESS. Configuramos este parámetro a su valor máximo en una PDA, mientras que la restante se configura al valor mínimo. Por tanto, la primera PDA es seleccionada como MPR. Medimos el tiempo de vida de la batería para cada PDA, manteniendo la luz de la pantalla encendida, en dos casos: en el primero, no se envía tráfico de datos, de modo que sólo aparecen mensajes OLSR a través de la red, y en el segundo, se envía un flujo UDP desde el portátil 1 hasta el portátil 2, a una tasa igual a la capacidad medida en este escenario.

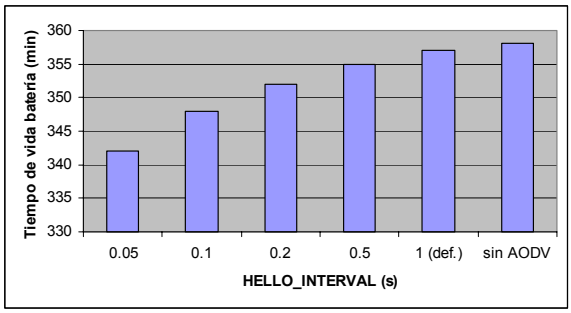


Fig. 5. Efecto de la configuración de HELLO_INTERVAL de AODV en el tiempo de vida de la batería en un escenario con topología en cadena de 2 saltos

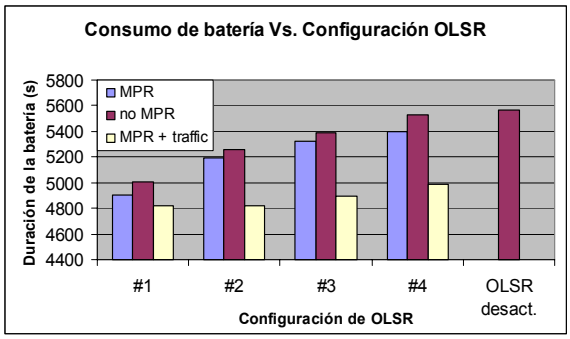


Fig. 6. Influencia de la configuración de parámetros de OLSR en el tiempo de vida de la batería de una PDA actuando como nodo retransmisor en un escenario con topología de 2 ramas y 2 saltos

Como se muestra en la Fig. 6, un incremento en la frecuencia de generación de mensajes de control provoca una reducción en el tiempo de vida de los dispositivos testeados. Un MPR consume más batería que un nodo ordinario debido a la generación de mensajes TC. Además, si un MPR actúa como retransmisor de datos, este efecto se enfatiza. Sin embargo, el impacto cuantitativo de la operación de OLSR en el consumo de batería no es particularmente significativo. Por ejemplo, el tiempo de vida de la batería de un MPR con la configuración de OLSR por defecto (es decir, configuración #3) es el 95'6 % de la lograda con OLSR deshabilitado. En nuestros experimentos, el peor de los casos se da utilizando la configuración #1, que implica un tiempo de vida del 86'5%, eso es, 12'5 minutos menos que deshabilitando OLSR y unos 10 minutos menos que usando la configuración de OLSR por defecto.

5.5 Observaciones

Los resultados obtenidos sugieren que usar configuraciones de AODV y OLSR con una frecuenca de generación de mensajes mayor puede ayudar a reducir la LCR, con un efecto mínimo en el tiempo de vida de la batería y, también en el ancho de banda disponible. Por ejemplo, reducir el HELLO_INTERVAL de AODV de 1 s a 200 ms decrece el ancho de banda disponible en un 10% en una topología en cadena de 4 saltos e introduce un consumo de batería adicional del 3%. Sin embargo, la misma configuración ayuda a reducir la LCR en un factor igual a 5 (es decir, de 0.2 a 0.4 s). En el caso de OLSR, protocolo menos adecuado para condiciones de alto grado de cambios en la topología debido a su naturaleza, se ha medido una reducción de unos 5 s (sobre los 13 s por defecto) en la LCR si se multiplica por 4 la tasa de envío de mensajes de control, con un consumo adicional de batería del 13.5% y de ancho de banda igual al 4%.

Por otra parte, debe tenerse en cuenta que condiciones de red distintas, tales como el tamaño de la red, el patrón de tráfico, los tipos de dispositivos y los patrones de movilidad, pueden determinar entornos distintos para los cuales las configuraciones de parámetros adecuadas pueden variar.

6. Conclusiones y trabajo futuro

La configuración de parámetros de los protocolos de encaminamiento en una red ad-hoc afecta al rendimiento de la misma en términos de ancho de banda disponible, latencia, reactividad frente a cambios en la topología y consumo de batería. Hemos analizado de forma teórica y empírica el impacto de los protocolos AODV y OLSR, su implementación y la configuración de algunos de sus parámetros más relevantes. Los resultados sugieren que, en los escenarios considerados, una configuración de parámetros distinta a la propuesta por defecto (tanto para AODV como para OLSR) puede conducir a una mejora significativa de la reactividad de los protocolos frente a cambios en la topología, sin incurrir en un consumo significativo de ancho de banda o batería. Sin embargo, debe tenerse en cuenta que la configuración adecuada dependerá de un conjunto de factores como el tamaño de la red, el patrón de tráfico, los tipos de dispositivos y los patrones de movilidad.

Como trabajo futuro, planeamos desarrollar mecanismos adaptativos de autoconfiguración de la frecuencia de envío de mensajes de control para los protocolos de encaminamiento analizados, que tengan en cuenta el citado conjunto de factores que determina las condiciones específicas de una determinada red ad-hoc.

Agradecimientos

Este trabajo ha sido financiado en parte por el FEDER y el Gobierno Español a través del proyecto TIC2003-01748.

Referencias

[1] C. Perkins, E. Belding-Royer, S. Das, "Ad hoc On Demand Distance Vector Routing (AODV)", RFC 3561, Julio 2003.

[2] T. Clausen, P. Jacquet, "Optimized Link State Routing Protocol (OLSR)", RFC 3626, Octubre 2003.

[3] I.D. Chakeres, E. Belding-Royer, "The Utility of Hello Messages for Determining Link Connectivity", International Symposium on Wireless Personal Multimedia Communications (WPMC) 2002.

[4] Implementación de AODV-Uppsala University: http://core.it.uu.se/AdHoc/ImplementationPortal

[5] Implementación de OLSR de UniK: http://www.unik.org

[6] H. Lundgren, E. Nordstrom, C. Tschudin, "Coping with Communication Gray Zones in IEEE 802.11b based Ad hoc Networks", WoWMoM'02, Septiembre 2002.

[7] Herramienta iperf: http://dast.nlanr.net/Projects/Iperf/

[8] Herramienta Ethereal: http://www.ethereal.com/

[9] I. Chakeres, E. Belding-Royer, "AODV Routing Protocol Implementation Design", International Conference on Distributed Computing Systems Workshops, 2004.

Diseño de una Red de Sensores Inalámbricos para Medida de Temperaturas en Ruedas y Ejes

R. Zubillaga[2], P. Bustamante[1,2], M. Aybar[2], J. Meléndez[1,2], N. Rodriguez[1], J. Ruiz[1]
Grupo de Telemática – Dpto. Electrónica y Comunicaciones
[1]CEIT y [2]TECNUN (Universidad de Navarra).
Manuel Lardizábal 15, 20018 – San Sebastián
Teléfono: 943 219 877 Fax: 943 311 442
E-mail: {rzubillaga, maaybar}@tecnun.es, {pbustamante, jmelendez, jruiz, nrodriguez}@ceit.es

Abstract. *Nowadays the temperature of bearings in high speed vehicles is determining their speed. It is necessary to know the value of this variable at each moment, in order to be able to prevent possible failures. Traditional solutions use wiring systems which are quite reliable but very expensive. This paper describes a new way of carrying out this measurement using a wireless technology system based on the "Industrial, Scientific and Medical" (ISM) band, becoming a safe and cheap solution. It describes how this wireless sensor measures and transmits information, as well as the laboratory tests used to obtain proper performance in an industrial environment. Although we only present the temperature sensor in this paper, the system can be expanded to the measurement of many other variables.*

1 Introducción

Conocer la temperatura a la que están trabajando distintos dispositivos ha sido siempre crucial en procesos industriales. Este es el caso de vehículos de alta velocidad en los que una temperatura excesiva en los elementos rodantes tiene como consecuencia una parada de emergencia para evitar posibles accidentes.

Las soluciones actuales recurren a dispositivos de toma de temperatura cableados, lo que supone un gasto añadido a la hora de la fabricación de los ejes y en el mantenimiento del vehículo. Estos dispositivos están duplicados o incluso triplicados para asegurar su correcto funcionamiento, lo que incrementa más su coste.

La red de Sensores de Temperatura Inalámbricos, en adelante STI, propuestos en este artículo, tienen como finalidad sustituir la funcionalidad de los actuales Sensores y a la vez reducir el gasto de instalación y mantenimiento de los mismos.

Para implementar este tipo de sensores, hay que tener en cuenta la hostilidad de un entorno de trabajo industrial, que determinará los elementos utilizados y los tiempos mínimos y máximos de refresco de la información a monitorizar, para poder estimar una vida de la batería adecuada a las políticas de mantenimiento del vehículo a sensorizar.

1.1 Estado del Arte

Los sensores inalámbricos aplicados a la industria se han hecho un hueco, sobretodo en la industria automovilística. Cada vez es mayor el grado de automatización de los vehículos y en consecuencia mayor el número de dispositivos electrónicos usados para el control del mismo. Se pueden clasificar en dos grupos: los destinados a aumentar el confort del usuario como pueden ser los sensores de lluvia, luminosidad y proximidad; y los destinados a la seguridad, como sensores de presión, temperatura, acelerómetros, etc. Estos son ejemplos de elementos que ayudan a conocer en todo momento el estado del vehículo. En el caso de los sensores de presión, que van dentro del neumático, no es posible utilizar ningún cableado, por lo cual nace la necesidad del sensor "*wireless*". Hay varios ejemplos de sensores de presión de la Rueda (TPM) comerciales, como los de Motorola [1], Sensonor [3] o Microchip [5] .

1.2 Necesidad de Implementación de un nuevo Sensor

La necesidad de implementación de un nuevo sensor de temperatura surge porque los sensores comercializados actualmente no cumplen los siguientes requisitos:

- Las políticas de tiempos de envío de la información y modos de configuración deben ser programables por el usuario final.
- El hardware del diseño del sensor debe estar adecuado a la mecánica de la ubicación final del mismo para facilitar su montaje y mantenimiento.
- Los sensores comerciales se enmarcan dentro de los sensores TPM, que monitorizan la temperatura y también presión y aceleración, que son variables no necesarias.
- EL alcance de los TPM actuales no es suficiente para la aplicación estudiada.

1.3 Medida de Temperatura en Rodamientos

Hoy en día el proceso de medida de temperatura en los rodamientos de las ruedas exige hacer un agujero en el eje para poder acceder la zona de medición o zona caliente. Este agujero además de ser de difícil acceso, lo que dificulta su instalación y mantenimiento, modifica la resistencia del eje y encarece sus costes de fabricación.

Además, debido a la complejidad del cableado que es necesario para alimentar y recoger la señal de los sensores, la electrónica debe situarse en el eje, donde está sometida a fuertes vibraciones, y debe estar provista de conectores suficientemente grandes para dar cabida al numeroso cableado

Los agujeros en el eje sirven de guía para el cableado del sensor de temperatura, desde la central de control de sensores hasta la zona caliente. En la Fig. 1 se puede ver el cableado necesario para traer la señal del sensor en los rodamientos hacia la electrónica de control, a través de los agujeros del eje.

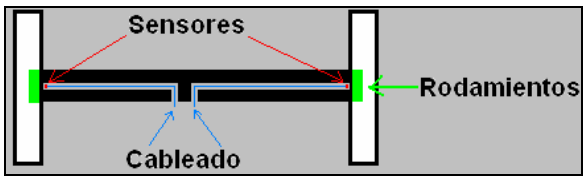


Fig. 1 Eje Cableado

1.4 Finalidad del Sensor

Todo la infraestructura utilizada actualmente conduce a que el mantenimiento del sistema de detección de zonas calientes sea excesivamente costoso en tiempo, recursos y dinero, especialmente si se tiene en cuenta que cada vehículo puede llevar más de 100 de estos sensores.

Con la introducción de los STI, sería posible colocarlos en lugares de fácil acceso, reduciendo de forma considerable los costes, la duración y la complejidad de las operaciones de mantenimiento. Así mismo, permitiría optimizar el diseño del eje y del equipo de monitorización de temperaturas de las zonas calientes, reduciendo los costes y la complejidad de su fabricación.

2 Especificaciones

La consecución de una red de Sensores lleva asociados unos requerimientos para el sensor de temperatura y para la centralita receptora de la información, derivados en su mayor parte del entorno de trabajo y de las necesidades de funcionamiento.

2.1 Sensor Inalámbrico de Temperatura

Se pueden diferenciar dos partes en el STI, el sensor de temperatura propiamente dicho y el módulo electrónico. Estos dos elementos tienen que cumplir unas especificaciones mínimas:

Sensor de Temperatura: Este elemento es una resistencia tipo PTC (Positive Temperature Coefficient) que tiene que ser capaz de medir temperaturas de entre 20ºC y 150ºC con una resolución de 1ºC como mínimo.

Módulo Electrónico: Este es el módulo que se encargará de muestrear la señal del sensor y enviarla por Radio a la central. A este conjunto de componentes se le exige que cumpla el rango de temperatura industrial de funcionamiento, que va desde los -40ºC hasta los 85ºC. Tiene que tener un identificador "ID" único para cada dispositivo, que además tiene que ser programable en el tiempo de instalación del sensor. Por otra parte tiene que guardar unos parámetros de configuración referidos a los intervalos de tiempo de envío de la temperatura, dados con el fin de ahorrar energía de la batería. Estos intervalos tienen unos valores por defecto, pero se tienen que poder cambiar junto con el ID en el momento del arranque. Los valores de tiempo de envío por defecto son especificados a continuación:

- T^a<40ºC: 128s. Si la zona caliente está a esta temperatura indica que el vehículo está parado. Es un intervalo de bajo riesgo.
- 40ºC≤T^a<60ºC: 64s. Este rango de temperatura no implica riesgo y se dará cuando el vehículo circule a baja velocidad, o será una zona de transición durante la puesta en marcha o la parada.
- 60ºC≤T^a<80ºC: 32s.: Este es el rango de funcionamiento normal durante la marcha del vehículo.
- T^a≥80ºC: 16s. Si la temperatura alcanza este rango habrá que prestar atención pues se considera zona de alto riesgo. Si lo hace durante periodos de tiempo prolongados habrá que contrastar con otro sensor o revisar el vehículo.

2.2 Central de Recepción de Datos

La central de recepción de datos, *centralita*, irá situada en una zona más protegida que la electrónica del STI, lo cual facilita su diseño. No obstante deberá cumplir requisitos de funcionamiento en rango de temperatura industrial entre -40ºC y 85ºC, ser capaz de alimentarse de la batería del vehículo con un rango entre 24v y 110v y su interface de comunicación con el resto de equipos ha de cumplir la norma RS-232, aunque se puede adaptar fácilmente a cualquier otro estándar, como RS-485 o CAN.

2.3 Requisitos RF

Para conseguir una reducción de costes mayor se ha pensado que la banda de trabajo del STI ha de ser una banda libre.

Una banda muy utilizada por equipos de radio frecuencia sin pago de licencia, es la banda ISM [1] (Industrial, Scientific & Medical). La ISM ofrece tres frecuencias básicas de funcionamiento: 434MHz, 868MHz y 2,4GHz. De estas es preferida la frecuencia de 868Mhz ya que la frecuencia de 434Mhz es muy usada por equipos PMR (Private Mobile Radio), lo cual supone un mayor riesgo a interferencias, y la frecuencia de 2,4Ghz tiene una mayor complejidad de diseño y sintonización.

3 Implementación

En este apartado se describen a modo funcional los distintos bloques que componen el STI y se adjunta su esquema eléctrico realizado en Orcad [8]. Para la realización del esquema eléctrico se han tomado como punto de partida los diseños de los fabricantes, tanto del microcontrolador como del *transceiver*.

En la siguiente figura se pueden ver los *Bloques funcionales del STI*:

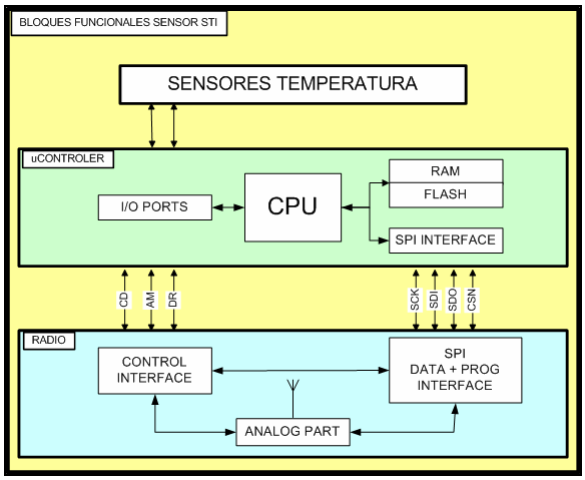


Fig. 2 Bloques Funcionales STI

Hay que destacar que el bloque del microcontrolador está formado por un PIC16LF876A (micro de bajo voltaje y bajo consumo) y el bloque de Radio por un nRF905.

El microcontrolador utiliza un convertidor A/D para leer el valor del sensor de temperatura. En la memoria *EEPROM* mantendrá grabados los parámetros de envío de la temperatura. A través del interface de comunicaciones SPI (Serial Port Interface) el microcontrolador se comunica con la Radio y la parámetros típicos de las radios. Utiliza tres líneas de control para poner la radio en un estado u otro: Recepción, Transmisión y modo Standby (bajo consumo).

En la siguiente figura se puede apreciar la foto de un sensor, con su antena para 868Mhz.

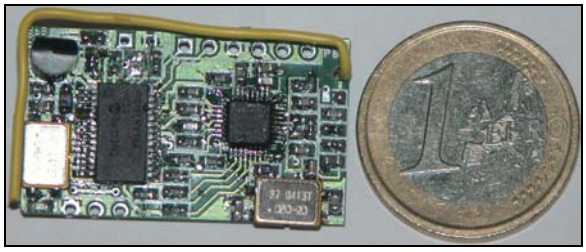


Fig. 3 Foto de un Sensor con su antena en 868MHz

En la *centralita*, los bloques funcionales son iguales con la diferencia de que ésta no tiene un sensor de temperatura, pero si tiene un interface serie RS232 de bajo consumo añadido, para poder comunicarse con el bus del vehículo.

La tecnología de fabricación de la centralita es de tipo mixto combinando SMD (Surface Mounting Devices) y TO (Through Hole), mientras que la tecnología del STI es enteramente SMD.

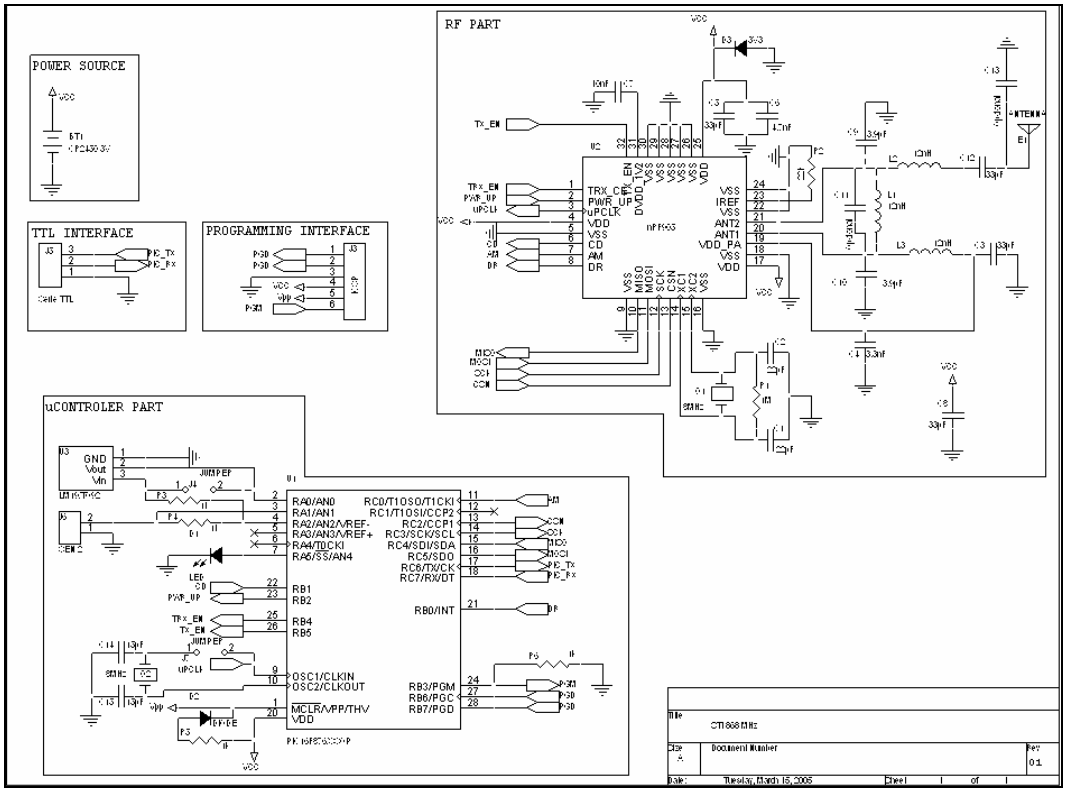


Fig. 4 Esquema Eléctrico STI

4 Comunicación Sensor Central

En este apartado se explica el protocolo de comunicaciones que se ha implementado entre la central y el sensor. Este protocolo se apoya en una característica del chip de RF que ha implementado el fabricante a nivel de hardware que se llama "*SockBurst Mode*". Se explica también el software creado bajo *Windows* para controlar y programar los STI's.

4.1 SockBurst

El nRF905 utiliza el modo "*Nordic Semiconductor ASA ShockBurst*™". El SockBurstTM, posibilita el uso de la alta velocidad de transmisión del nRF905 sin la necesidad de usar un microcontrolador (MCU) de alta velocidad en el procesado de señal, minimizando el coste del desarrollo. El nRF905 tiene integrados todos los elementos de alta velocidad referidos para el protocolo de RF y dispone de un interface SPI, cuya velocidad gestiona el MCU para pasar la información entre el chip de radio y el MCU. De esta forma se consigue que la parte digital del diseño trabaje a baja velocidad, mientras se maximiza la velocidad de comunicación en el Link RF, reduciendo de este modo el consumo medio de corriente en la aplicación, necesario para cumplir el requisito de bajo consumo.

El funcionamiento del SockBurstTM se puede separar en dos modos, el modo de recepción y el de transmisión:

- En modo Recepción, las señales de *Address Match* (AM) y *Data Ready* (DR) que la radio posee, informan al MCU cuando una dirección y un Payload válidos han sido recibidos respectivamente.

- En modo transmisión, el nRF905 genera automáticamente el preámbulo y el CRC y la señal *Data Ready* (DR) informa al MCU que el dato ha sido enviado. De este modo se reducen costos en el MCU por el menor número de memoria a usar y por el tiempo de desarrollo de software.

4.2 Composición de Tramas

En la Fig. 5 se muestran cómo están compuestas las tramas enviadas desde los distintos STI's hacia la centralita. La trama está compuesta por un total de 17 bytes, que son los que viajan por aire.

En dichas tramas, la MCU solo tiene que colocar la información que quiere enviar en los registros apropiados de la radio, a través del SPI, y dar la señal de envío. Esos registros son el "TARGET ID", que indica el dispositivo de destino y el "PAYLOAD" que es en la que se envía la información de los sensores. El Preámbulo y CRC son generados automáticamente por la radio, como ya se ha comentado anteriormente

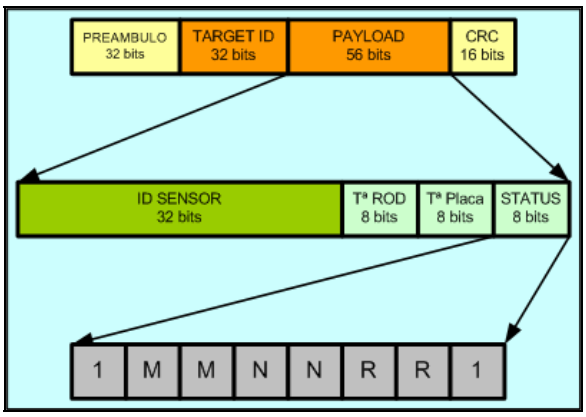


Fig. 5 Trama ShockBurst

El PAYLOAD tiene una longitud de 56 bits y el byte de STATUS, está compuesto como sigue:

MM: dos bits que indican el modo de trabajo o la Frecuencia de Transmisión de las medidas y depende de la temperatura.

NN: dos bits que indican el número de sensores instalados en el STI. Variará entre 1 y 2.

RR: son dos bits para futuras implementaciones

4.3 Funcionamiento

Teniendo en cuenta que no habrá ninguna actuación externa al sensor durante todo su funcionamiento, a excepción del momento del arranque, se ha pensado en un sistema que permita programarlo en el instante de arrancar.

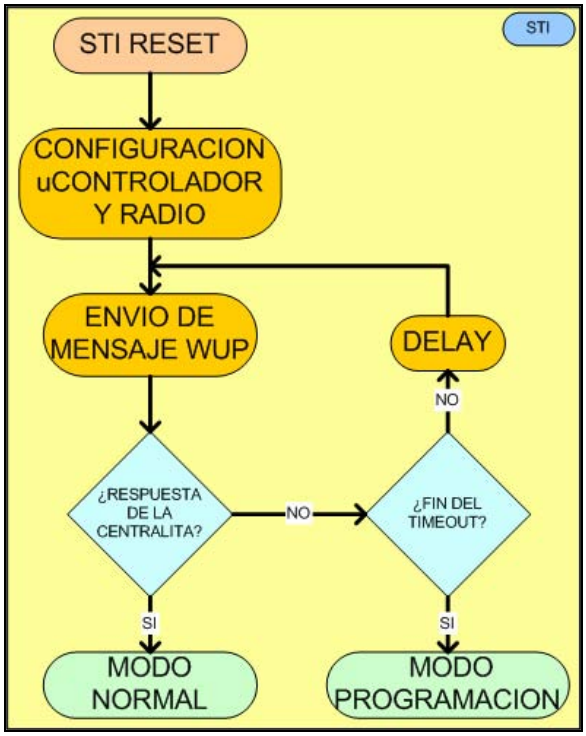


Fig. 6 Diagrama de Flujo del Sensor

Como se aprecia en la Fig. 6, una vez inicializado el sensor y durante un tiempo preestablecido, éste enviará un mensaje de "*wake up*", WUP, por aire. La *centralita,* que estará a la escucha, esperando para configurarle, le contesta al sensor para que éste entre en *modo programación*, y así se le podrán cambiar los parámetros de configuración de los tiempos de envío e incluso asignarle un nuevo ID. En caso contrario, y pasado un tiempo, pasará al modo de *funcionamiento normal* con los parámetros cargados por defecto, es decir, toma la temperatura y la envía por aire, según la política de tiempos establecida.

A continuación se explican los dos modos de funcionamiento:

Modo Configuración: Es posible establecer una comunicación semi-duplex entre la STI y la central para poder programar ciertos parámetros del STI. Este modo solo es accesible durante unos 30 segundos después de *resetear* el STI. Pasado este tiempo, si la central no ha enviado ningún comando al STI, éste pasará automáticamente al modo de trabajo, usando como parámetros unos valores por defecto.

Modo Normal: En este modo el sensor no tiene activa su parte Rx, por lo que no es posible establecer una comunicación con él. Su funcionamiento se basa en generar y enviar la trama de PAYLOAD siguiendo la política de tiempos expresada por los parámetros de configuración, que será la política usada por defecto. Para evitar posibles colisiones cuando el STI o la central envían su trama, se ha establecido una escucha de canal durante 5ms. Si durante ese tiempo el canal permanece libre, es decir no hay portadora, se procede a la transmisión. Si en ese tiempo se produce alguna transmisión, el STI que quiere enviar esperara un tiempo aleatorio que depende directamente de su ID. Pasado ese tiempo, vuelve a intentar la transmisión.

4.4 SOFTWARE

Se ha diseñado un software que permite conectar una centralita STI y gestionar una red de 4 STI de forma sencilla. El software consta de dos ventanas de diálogos, el diálogo principal y el diálogo de configuración.

Diálogo principal: Llamado "Sensores de Temperatura", consta de unos hitos en los que se puede ver el estado de cada sensor, su ID, y la temperatura que está midiendo. Se puede comprobar de forma rápida y sencilla todos los datos que interesan de cada sensor, como son su ID y la temperatura que está registrando. También informa mediante unos testigos luminosos del estado del canal serie, si está en transmisión o en recepción.

Esta ventana contiene un botón "configuración" desde el que se accede al siguiente diálogo.

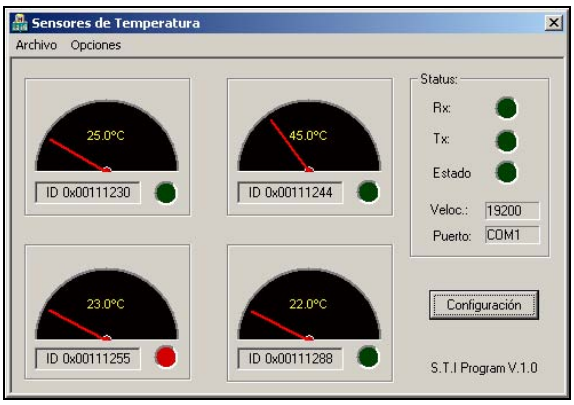


Fig. 7 Diálogo Principal

Diálogo Configuración: Desde este diálogo podremos cambiar el ID del STI, así como la base de tiempos. Este diálogo se podrá llamar siempre que tengamos un nuevo STI que haya enviado un mensaje indicando que está en ese modo. Seleccionados los parámetros necesarios se pulsará el botón "Enviar" y si la transmisión es correcta se indicará mediante un mensaje en pantalla.

Fig. 8 Diálogo de Configuración

5 Validación del Sistema

Para verificar el correcto comportamiento de la red, se ha sometido a los distintos prototipos de los STI's a una serie de pruebas para comprobar que cumplen las siguientes especificaciones: Alcance de Transmisión al aire libre y en entorno hostil, tiempos de los distintos estados de funcionamiento y descarga de baterías. Los resultados de estos análisis se recogen a continuación.

5.1 Alcance de la transmisión

Las pruebas de alcance se han realizado en 3 entornos distintos: el primero al aire libre, el segundo en una zona despejada pero dentro de una nave industrial, y el tercero usando un vehículo de alto tonelaje como campo de evaluación y situando la *centralita* en una posición intermedia de la longitud total del vehículo y el sensor en distintas posiciones alrededor de los ejes de medida.

En todas las pruebas se han utilizado cuatro sensores STI en 868MHz y la Centralita conectada a una ordenador portátil, con el programa que ya se describió en el apartado anterior. Los sensores se han colocado en un encapsulado o caja de resina, similar a la que lo albergará realmente. La función de esta caja es la de protección del STI frente a humedades, polvo, vibraciones y posibles golpes.

Cada sensor STI realiza tres envíos consecutivos, de forma aleatoria y con la misma trama, durante un tiempo de 20ms. Dichos datos se reciben en la centralita correctamente, mostrando en el programa los valores de la temperatura y los ID's de cada uno de los sensores.

Zona al Aire libre: Las premisas de la prueba son un día despejado , con temperatura ambiente de 12ºC. Se realiza la prueba alcanzándose una longitud máxima de 90 metros.

Zona nave Industrial: Las premisas de la prueba son temperatura ambiente de 20ºC y una nave de 100 metros largo por 3 de ancho y 4 de alto y con puentes grúa trabajando. Se consigue una longitud máxima de 80 metros.

Zona Vehículo Pesado: Las condiciones iniciales son una temperatura ambiente 15ºC y una nave industrial con ruidos electromagnéticos provenientes de puentes grúa y otros elementos de radio control, situados en un área de 80 metros de diámetro aproximadamente. Las posiciones de la Centralita y el STI son las mostradas en la figura:

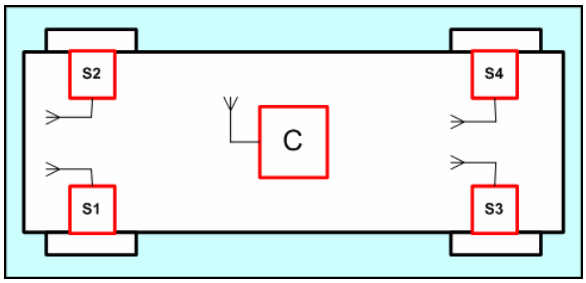


Fig. 9 Vehículo de Pruebas

Se consigue que en todos los puntos de la prueba la señal sea recibida.

5.2 Tiempos de funcionamiento

El STI tiene cuatro modos de funcionamiento en su ciclo:

- *Modo 1:* Tanto el microcontrolador como la radio están en modo "Power Down", o modo de bajo consumo de energía. Es el estado en el que se pasan la mayor parte del tiempo y dura 2,45s cada vez. En este modo el consumo es de 3uA.

- *Modo 2:* Es un tiempo de transición y es lo que tarda el microcontrolador en pasar del modo de bajo consumo a estar totalmente operativo, dura menos de 5ms.

- *Modo 3:* Es el tiempo de medida de la temperatura a través del convertidor analógico digital. Su duración es de menos de 1ms.

- *Modo 4:* Es el tiempo dedicado a la carga de los registros de la radio y a la transmisión. Es crucial que sea lo menor posible, ya que en este modo el consumo del STI es elevado. Su duración es de 19,5ms y el consumo es de 23mA.

Estos modos de funcionamiento, obtenidos experimentalmente, son los tiempos base de los que se tomarán las referencias para los 2 ciclos de funcionamiento. El ciclo de funcionamiento A está compuesto por los modos 1, 2 y 3 y el ciclo de funcionamiento B está compuesto por los 4 modos. De acuerdo con la política de tiempos del STI, un ciclo de funcionamiento completo está formado por ciclos A y B tal como se muestra en la siguiente imagen:

Fig. 10 Ciclos de Funcionamiento

La variable "n", que es el número de ciclos de tipo A, dependerá de la temperatura medida en ese momento.

Esto sirve para hacer una estimación del consumo del STI durante un funcionamiento real.

5.3 Descarga controlada de baterías

Para que el sistema tenga una autonomía mínima de de 2 años se han supuesto los tiempos de funcionamiento siguientes, que son la situación más desfavorable posible en un caso real.

- Si T^a>80ºC se hará un envío cada 16s. Se supone para la prueba que este estado ocupa el 50% del tiempo de funcionamiento total.

- Si 80ºC>T^a>60ºC se hará un envío cada 32s, ocupando un total del 8% del tiempo de trabajo.

- Si 60ºC>T^a>40ºC los envíos se realizarán cada 64s, siendo un 8% del total del tiempo de trabajo.

- Si T^a<40ºC el vehículo estará parado y el envío se realizará cada 128s, ocupando un 34% del tiempo total.

Realizando medidas de la energía consumida en cada modo reflejado en la política de envíos, y aplicando esta lógica de estados, se estima que el consumo medio del sensor está en 21 microamperios. Esto implica que necesitamos una batería capaz de entregarnos en dos años 370 mA de carga. Esto se justifica en la siguiente expresión:

Carga = 0.021mA*(2*365*24)h ~=370mA

Tomando como referencia una pila del tipo Li-Ion CR2450 de Panasonic de 3V y 650mAh, teóricamente se puede conseguir este funcionamiento.

Teniendo en cuenta que el rango de tensión de alimentación que se puede aprovechar está comprendido entre 2voltios, que es el límite de funcionamiento del STI y los 3 voltios que da la pila inicialmente, se somete a esta batería a una descarga como la siguiente:

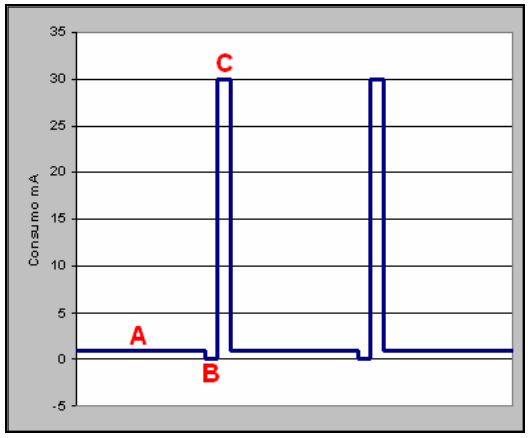


Fig. 11 Descarga Batería

Se pretende con esto caracterizar la curva de descarga de la pila:

- Zona A: Se le somete a un consumo continuo muy pequeño, que para la prueba se sitúa en 1mA durante el 97% del tiempo. Simula el modo de bajo consumo del micro y la radio.

- Zona B: Es una zona de recuperación.

- Zona C: Se caracteriza porque son pulsos de muy corta duración, de 20ms pero de gran exigencia de carga, unos 30mA. Simula el momento de la transmisión.

Estos datos se van registrando tomando el valor de la tensión en los bornes de la batería con un convertidor AD. Esta prueba se extiende durante varios días hasta que la tensión en bornes de la batería alcanza los 2V, unas 400 horas después. Así la energía entregada por la misma, siendo la temperatura ambiente de 15ºC, ha sido la siguiente:

Econtinua = 400h*0.971*1mA ~= 388mAh

Epicos = 400*0.00971*30mA ~= 116mAh

Etotal = 388 + 116 ~=500mAh

El rendimiento obtenido de la pila en estas condiciones es del 77% aproximadamente.

A la vista de estos datos y en las condiciones especificadas se ve que es posible cumplir los requerimientos de vida útil del STI.

Conclusiones

Los sensores STI han sido diseñados para facilitar la instalación y mantenimiento de los sistemas de control de temperatura en lugares de difícil acceso, como son los ejes y rodamientos de vehículos pesados de una forma rápida y sencilla. Sus características más relevantes son:

- Disponer de un sistema de telemetría, basado en una red de sensores, de bajo costo y bajo consumo de energía.
- Lucha por acceso al medio, permitiendo que, aunque haya varios sistemas juntos en una sola frecuencia, siempre logren enviar la información a la *centralita*.
- Facilitar la instalación de este tipo de sensores sin hacerlos dependientes de cableados ni mantenimiento.
- Reducir los costos de fabricación de los elementos sensorizados, al no tener que modificar el proceso de fabricación, como en el caso de ejes.
- Posibilidad, mediante un pequeño cambio de hardware del STI, de convertirlo en un Sensor para medir cualquier otra variable física o química, como aceleración, presión y concentración de distintos gases.

Antes era necesario una parada técnica de varios días de los vehículos para poder sustituir el sensor dañado y su cableado, lo que suponía pérdidas económicas importantes debidas a que el vehículo no está en circulación y a los costos añadidos del mantenimiento para el que era necesario varias personas. Ahora este tiempo de parada es reducido enormemente al no tener que sustituir el cableado que suponía desmontar el vehículo.

El STI se ha desarrollado usando tecnologías SMD para reducir el tamaño del mismo, convirtiéndolo en un dispositivo de tamaño y prestaciones único.

Referencias

[1] Bandas ISM definidas por la ITU-T en S5.138 y S5.150. www.itu.int/ITU-T/.

[2] TPM system SG2011 de Freescale (Motorola). www.freescale.com.

[3] TPM Sensonor SP12 www.sensonor.com.

[4] Corporate Headquarters Microchip Technology Inc. 2355 West Chandler Blvd. Chandler, Arizona, USA www.microchip.com

[5] TPM Sensor de Microchip. Nota de aplicación AN238. www.microchip.com.

[6] Nordic Semiconductor ASA Trondheim HQ Vestre Rosten 81 7075 Tiller Norway. www.nordicsemi.no.

[7] AEL Crystals Limited Module D Airtech 2 Jenner Rd Flemming Way Crawley West Sussex RH10 2GA www.aelcrystals.co.uk

[8] ORCAD www.orcad.com

[9] Panasonic http://www.panasonic-industrial.com

Estudio de MANETs híbridas con gateways móviles

Alicia Triviño Cabrera, Eduardo Casilari Pérez
Departamento de Tecnología Electrónica. Universidad de Málaga.
ETSI de Telecomunicación. Campus de Teatinos.
29071 – Málaga
Teléfono: 952 13 71 91 Fax: 952 13 21 16
E-mail: atc@dte.uma.es, ecasilari@uma.es

Abstract*. Mobile Ad hoc NETworks (MANET) were conceived to satisfy the requirements of communication among mobile devices avoiding the use of deployed infrastructures. Due to the importance the Internet has acquired, this initial concept is being extended. By the use of an access router it is possible to achieve global connectivity in the boundaries of the MANET, i.e. devices that are located inside the coverage of the access router can use it directly. For the rest of the nodes, a gateway is required in order to route the packets as well as to provide an appropriate network prefix. This paper explains the behaviour of the support for connecting the MANET to the Internet based on mobile multi-gateway.*

1 Introducción

Las redes móviles ad hoc (MANETs) surgen con el propósito de establecer la comunicación entre terminales inalámbricos sin la necesidad de recurrir a elementos centralizados dedicados a la gestión de los recursos radio, al encaminamiento o a la conmutación. Inicialmente fueron ideadas para aquellas situaciones en las que no es posible acceder a las redes de telecomunicaciones tradicionales debido a catástrofes naturales o a eventos bélicos. Entornos en los que la esporadicidad de la red desaconseje el gasto asociado a su implantación también pueden beneficiarse del uso de redes ad hoc.

Al igual que ocurre en las WLANs (*Wíreless Local Area Networks*) si un terminal desea enviar paquetes a otro dispositivo que se encuentra en su zona de cobertura es posible comunicarse directamente con él. No obstante, si esta condición no se verifica se necesitará la cooperación de nodos intermedios que reenvíen el paquete de datos hasta un terminal que sí se sitúe en la cobertura del destinatario.

La elección de los elementos intermedios así como la metodología empleada para conocerlos diferencia los protocolos de encaminamiento elaborados. En esta línea se han propuesto múltiples protocolos: AODV (*Ad hoc On-demand Distance Vector*), DSR (*Dynamic Source Routing*), OLSR (*Optimized Link State Routing*), TBRPF (*Topology Dissemination Based on Reverse Path Forwarding*), DYMO (Dynamic MANET on Demand), etc. Todos ellos resuelven la comunicación interna, es decir, entre terminales pertenecientes a una misma MANET.

Sin embargo si se desea proporcionar acceso a Internet a las redes ad hoc, convirtiéndolas en redes híbridas, es preciso añadir una serie de mecanismos a cualquiera de los protocolos expuestos anteriormente. La incorporación de un *router* de acceso a la red ad hoc es inevitable ya que se requiere de un elemento que sirva de enlace con una red accesible a Internet. Debido al encaminamiento jerárquico que domina Internet, el *router* de acceso proporciona el prefijo de red que debe ser empleado por aquellos dispositivos que deseen comunicarse con Internet a través de él. El *router* de acceso envía esta información periódicamente a través de mensajes de aviso de *router* tal y como se describe en el protocolo NDP (*Neighbour Discovery Protocol*) de IPv6 [1]. Es importante destacar que la normativa de NDP especifica que los mensajes defenidos en ella no pueden ser reenvíados. Para la red MANET ésta es, sin lugar a dudas, una limitación crítica pues sólo los dispositivos que se encuentran a un único salto del router de acceso podrían conocer el prefijo de red a emplear. Este es uno de los motivos de la exigencia de un elemento adicional, denominado *gateway*, encargado de propagar por toda la red esta información. Una vez recibido el prefijo los terminales de la red inician el proceso de autoconfiguración de direcciones para obtener una dirección IP global [2].

La Fig. 1 muestra los requisitos a considerar en la arquitectura de una red ad hoc conectada a Internet. De manera genérica se ha excluido el *gateway* de la MANET, tal y como se definió en los primeros mecanismos que aparecieron respecto a este tema (Wakikawa, Jelger). Otros mecanismos proponen una simplificación y autoconfiguran como *gateway* a uno de los dispositivos componente de la MANET.

En este artículo se describen los diferentes mecanismos publicados que proporcionan acceso a Internet a una MANET, centrándose en la explicación del procedimiento basado en *gateways* móviles. A su vez se muestra el comportamiento de dicho procedimiento a través de los resultados de las simulaciones realizadas.

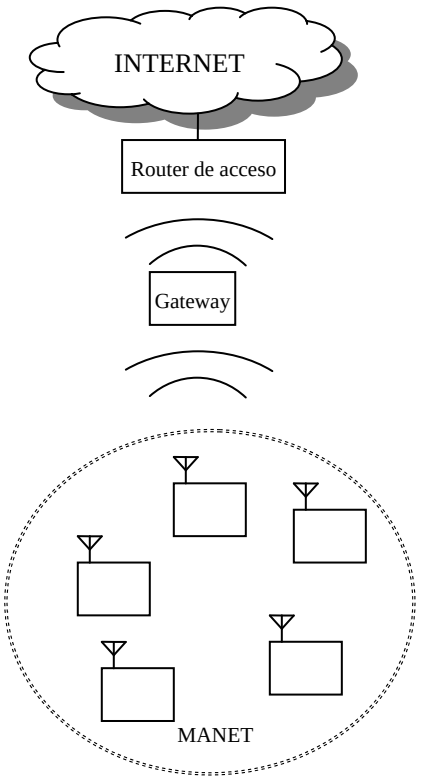


Figura 1. Arquitectura genérica de conexión a Internet en MANET.

2 Descubrimiento del *gateway*

Se han desarrollado diversas técnicas para lograr la conectividad a Internet en redes ad hoc. Generalmente se suele usar el paradigma de IPv6 otorgando a los terminales de la MANET una dirección IPv6 global [3]. Este motivo explica que habitualmente los procedimientos se califiquen como de conectividad global.

A continuación se detallan los mecanismos propuestos que mayor relevancia han alcanzado dentro de la comunidad investigadora. Para su explicación se va a seguir la terminología expuesta en NDP.

- **Conectividad global propuesta por Wakikawa** [4]. Se basa en incorporar al *router* de acceso un *gateway* fijo dedicado al encaminamiento ad hoc y propagación del prefijo de red. Permite diversos tipos de comportamiento asociados al *gateway*. Por un lado se encuentra la concepción proactiva en la que el *gateway* envía periódicamente mensajes de *broadcast* por toda la red donde se incluye el prefijo de red del router de acceso. Este tipo de paquetes se denominan mensajes de aviso de *router* modificados o MRA (*Modified Router Advertisement*).

  Por otro lado, en la percepción reactiva los terminales que desean iniciar una comunicación con Internet piden el prefijo a través de un mensaje de solicitud de *router* modificado o MRS (*Modified Router Solicitation*). Recibido este mensaje, el gateway responde con un MRA *unicast*.

  También existe el planteamiento híbrido. En este escenario el *gateway* manda MRA periódicamente con un tiempo de vida limitado a un número de saltos que no abarca toda la red. Los nodos que no reciben estos mensajes procederán como en la técnica reactiva para obtener la conexión a Internet.

- **Continuidad de prefijo** [5]. Pretende ser una mejora del presentado por Wakikawa al intentar reducir la carga en la red. Esta propuesta de Jelger *et al.* se centra en un entorno donde coexistan múltiples *routers* de acceso, cada uno de ellos con un *gateway* fijo asociado. Los nodos que reciben mensajes MRA de diferentes *gateways* eligen el *router* de acceso al que desean conectarse y sólo retransmiten el mensaje del *router* elegido. Debido a esta característica los terminales de la red están seguros de que existe una ruta hacia el *router* de acceso que utiliza el mismo prefijo de red. Es por ello que recibe la denominación de continuidad de prefijo.

- **Múltiples *gateways* móviles** [6]. Intenta minimizar los requisitos para la conexión a Internet en redes ad hoc y pretende equipararla a las exigencias de una red WLAN. También suele denominarse como mecanismo basado en múltiples gateways móviles. Es la estrategia analizada en este artículo.

3 Mecanismo con *gateways* móviles

Los *routers* de acceso (AR) son incapaces de encaminar un paquete destinado a un terminal que no se encuentre en su zona de cobertura. Esta limitación se debe a la ausencia de implementación de protocolos ad hoc en el *router* de acceso. Los mecanismos propuestos por Wakikawa y por Jelger solventan esta limitación asociando a cada *router* de acceso un dispositivo fijo localizado dentro del área de cobertura del *router*. Este terminal es un *gateway* IPv6/MANET. Todo paquete destinado a un terminal de la red ad hoc es enviado desde el AR hasta el *gateway* y éste lo encamina siguiendo un protocolo ad hoc. De igual manera, cuando un nodo móvil desea establecer una comunicación con Internet, envía sus paquetes a través del *gateway* que reenvía los paquetes directamente al *router* de acceso.

Esta exigencia de acompañar al *router* de acceso con un dispositivo adicional complica la infraestructura y es contraria al objetivo de autosuficiencia de las redes ad hoc: proporcionar acceso a una red WLAN difiere

de proveer conectividad global en una red ad hoc. Con el propósito de solventar este condicionante un grupo de investigación de Samsung desarrolló un nuevo mecanismo basado en la configuración dinámica de nodos móviles como *gateways* [6].

La idea fundamental reside en el hecho de que todos los terminales que componen la red ad hoc poseen implementados tanto IPv6 como un protocolo ad hoc. Por lo tanto, cualquier dispositivo que se encuentre a un único salto del *router* de acceso puede realizar las tareas asociadas al *gateway* fijo de las propuestas de Wakikawa o de Jelger.

Es importante señalar que no se impone ninguna restricción en el movimiento del dispositivo elegido para actuar como *gateway*. Sería posible, pues, que dicho dispositivo saliese de la zona de cobertura del *router* de acceso por lo que es preciso emplear un método para que alguno de los dispositivos que sí se encuentren a un salto se autoconfigure como nuevo *gateway*.

3.1 *Gateways*

Tal y como se ha mencionado anteriormente, cualquier dispositivo que se encuentre a un único salto del *router* de acceso puede actuar como *gateway*. Uno de estos va a ser elegido para actuar como tal. Es el denominado *gateway* por defecto y sus funciones son las siguientes:

- Encamina los paquetes procedentes de Internet a un terminal de la MANET. Debido al direccionamiento jerárquico que caracteriza IP, el *router* de acceso detecta con facilidad que el paquete va destinado a un dispositivo de la red ad hoc. Al carecer el *router* de acceso de mecanismos de descubrimiento de rutas en redes ad hoc, éste se lo reenvía al *gateway*.

- Encamina los paquetes procedentes de la red ad hoc hacia Internet. El gateway por defecto actúa, por lo tanto, como conexión a Internet.

- Propaga el prefijo de red para que los dispositivos autoconfiguren su dirección IP. El *router* de acceso envía periódicamente un mensaje de aviso de *router* o RA (*Router Advertisement*) que contiene el prefijo de red a emplear. Al pertenecer estos mensajes al protocolo NDP presentan la restricción de no poder ser reenviados [1]. Esta particularidad impide la autoconfiguración de direcciones en terminales alejados en más de un salto. Es necesario, pues, introducir modificaciones para que todos los nodos de la red ad hoc puedan recibir el prefijo de red.

Para lograr completar sus funcionalidades el *gateway* por defecto manda periódicamente mensajes modificados de aviso de *router* o MRA (*Modified Router Advertisement*). Son equivalentes a los RA, salvo que se permite su reenvío, por lo que alcanzaría todos los nodos de la red ad hoc.

Los terminales que reciben estos paquetes conocen, por un lado, si continúan asociados al mismo *router* de acceso ya que fácilmente pueden comparar el prefijo que les llega contenido en el mensaje con el de su dirección IP. En el caso de que carezcan de dirección IP global podrían iniciar la autoconfiguración [2]. Adicionalmente, este tipo de mensajes les otorga la información necesaria para construir una ruta hacia el *gateway*, elemento al que deben enviar sus paquetes dirigidos a Internet.

Dentro del área de cobertura del *router* de acceso pueden encontrarse más de un terminal móvil. Sólo uno de ellos se configura como *gateway* por defecto. El resto de los nodos se denominan *gateways* candidatos. Estos elementos pueden ser empleados en tareas adicionales para distribuir el tráfico tal y como se explicará posteriormente.

Aquellos terminales que deseen comunicarse con Internet y carezcan de información actualizada del *gateway* pueden iniciar una solicitud de *router* modificado o MRS (*Modified Router Solicitation*). Tras la recepción de este mensaje *broadcast*, el *gateway* por defecto responde con un MRA *unicast*. Se puede establecer la opción de que terminales intermedios proporcionen la información que almacenan sobre el *gateway* por defecto.

La Fig. 2 representa la arquitectura asociada a este mecanismo. En ella se aprecia que en la zona de cobertura del *router* de acceso coexisten un *gateway* por defecto (DG) y dos *gateways* candidatos (CG). Estos tres elementos son componentes de la MANET.

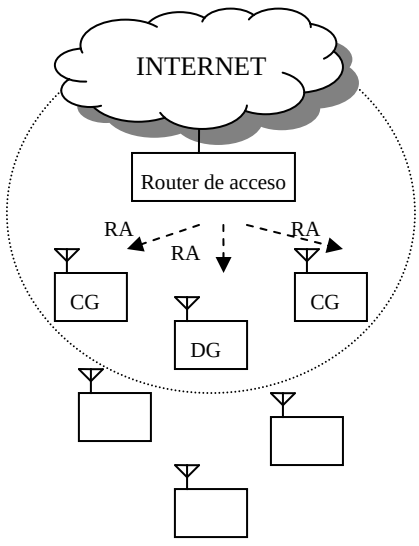


Figura 2. Arquitectura de Singh para conexión a Internet en MANET.

3.2 Elección del gateway por defecto.

Los *gateways* candidatos esperan recibir periódicamente mensajes MRA. Cuando detectan que esto no sucede entienden que el *gateway* por defecto anterior ha dejado de operar como tal, bien sea porque salió de la zona de cobertura del *router* de acceso o bien porque se desconectó.

Tras un tiempo de guarda desde el tiempo previsto para la recepción del MRA, los *gateways* candidatos arrancan un temporizador con un valor aleatorio cuyo valor máximo coincide para todos ellos. Una vez agotado el temporizador, el gateway candidato comprueba si se ha recibido un nuevo MRA en dicho intervalo. Si así hubiese ocurrido, ya habrá configurado como *gateway* por defecto el origen de dicho mensaje. En caso contrario se autoconfigura como *gateway* por defecto enviando inmediatamente su paquete de MRA para que el resto de la red detecte sus nuevas funcionalidades.

Aunque en el borrador del mecanismo no se indicaba, los autores de este artículo han apreciado que en algunas ocasiones la diferencia de los valores de los temporizadores no equivale al retardo de la propagación del mensaje MRA a través de varios saltos. En estos casos múltiples *gateways* candidatos se autoconfiguran como *gateway* por defecto. Estando expuesta claramente en dicho borrador la condición de que sólo puede existir un *gateway* por defecto para cada *router* de acceso, se ha percibido la necesidad de añadir una técnica para garantizar esta restricción. El mecanismo es bastante sencillo: cualquier *gateway* por defecto que reciba algún mensajes MRA cuyo origen posea una dirección IP menor que la suya se desconfigura automáticamente como *gateway* por defecto. Así pues el *gateway* con dirección IP menor prevalece frente a otros.

3.3 Distribución del tráfico

Con lo explicado hasta ahora la comunicación entre terminal móvil y terminal fijo requiere el uso del *gateway* por defecto. Sin lugar a dudas esta característica puede implicar un cuello de botella en el sistema. Una posible solución consiste en distribuir el tráfico por los *gateways* candidatos.

Basándose en IPv6 es posible exigir que un paquete circule por una serie determinada de terminales siempre y cuando se especifiquen correctamente en la cabecera adicional de encaminamiento de *proxy* de dicho protocolo [7].

Aquellas fuentes que deseen emplear estos dispositivos preguntarán periódicamente por todos los *gateways* a través del envío de mensajes modificados de *router* o MRS y estos responderán con los correspondientes mensajes modificados de aviso de *router*.

Pueden aplicarse diversos criterios para elegir, de las respuestas, el *gateway*, candidato o por defecto, más adecuado. Una primera opción sería elegir el que se encuentre a menor número de saltos de la fuente con el propósito de reducir el retardo. Esta idea es bastante simple de computar e incluso permite una optimización. Como todos los nodos de la red conocen el número de saltos de la ruta asociada al *gateway* por defecto, es posible reducir el número de respuestas si en la petición que realiza el terminal origen se fija un tiempo de vida equivalente a dicha distancia. De esta manera los *gateways* candidatos que se encuentren más alejados respecto al *gateway* por defecto no recibirán la petición.

Otra opción consiste en basarse en el tráfico que están soportando los *gateways* y seleccionar el menos cargado. Para ello los *gateways* deben contabilizar el tráfico que están enviando y recibiendo. Esta información se incluirá en los mensajes modificados de *router* dentro del campo de opciones. El nodo origen elegirá aquel que posea un valor menor.

4 Simulaciones

El análisis del mecanismo basado en *gateways* móviles ha precisado del desarrollo de un módulo software a partir del trabajo publicado por Alex Hamidian [8]. Este módulo se ha integrado en la herramienta *Network Simulator*, ns-2.1.9b [9] bajo entorno Linux.

Los escenarios probados se caracterizan por poseer un área de simulación de 1500 m x 300 m, situándose el *router* de acceso en el centro. 50 nodos componen la red ad hoc moviéndose según el conocido y extendido modelo de *Random WayPoint*. Para evaluar distintos escenarios de movilidad se modificó la velocidad máxima y el tiempo de pausa, tal y como se indica en la tabla 1. Dos variaciones significativas han sido incluidas en el patrón de movimiento respecto a la empleada en la mayoría de la bibliografía relacionada. Siguiendo la recomendación expuesta por Yoon *et al*. se ha fijado una velocidad mínima que garantiza una mayor estabilidad de los resultados [10]. Por otra parte, el tiempo de pausa se ha correlado con la esperanza matemática de la velocidad [13].

El tráfico generado se asocia a diez fuentes CBR que mandan paquetes de 512 Bytes a un nodo en Internet, empleándose por lo tanto el *router* de acceso. El resto de los parámetros de la simulación quedan recogidos en la tabla 1.

Como protocolo de encaminamiento dentro de la red ad hoc se ha escogido AODV por su amplia difusión [11]. Se trata de un protocolo reactivo que permite el descubrimiento de rutas así como la detección de la ruptura de las mismas. Dentro de las opciones que permite dicho protocolo se ha escogido la reparación local de las rutas en caso de ruptura.

Tabla 1. Parámetros de las simulaciones realizadas

Área de simulación	1500 m x 300 m
Nº de terminales	50
Patrón de movilidad	Veloc. máx = [1.1,5,10,20] m/s. Velocidad mínima = 1 m/s Pausa = [0%, 50%]
Modelo de tráfico	10 fuentes CBR Tasa = 4 paquetes/s Tamaño paquete = 512 B
Tiempo simulación	5000 s
Rango de transmisión Inalámbrica de nodos	250 m
Nº de ejecuciones	3
Protocolo ad hoc	AODV
Nivel de enlace	802.11a RTS/CTS habilitado Tasa de Datos = 2 Mbps
Tamaño cola interna	64 paquetes

Las prestaciones del mecanismo se han cuantificado con los siguientes parámetros:

- PDR (*Packet Delivery Ratio*). Se define como el cociente entre los paquetes de datos recibidos por el nodo destino, que se encuentra en la red fija accesible a través del router de acceso, y los enviados por los terminales móviles pertenecientes a la MANET.

- Retardo medio. Representa la media de los retardos extremo a extremo que sufren los paquetes de datos. Este parámetro incluye el tiempo que permanece en las colas de los nodos, retransmisiones en el nivel MAC y reenvío a través de múltiples nodos.

- *Overhead* normalizado. El *overhead* equivale a la suma de paquetes que no son de datos utilizados para proporcionar encaminamiento ad hoc y acceso a Internet En este sentido cada reenvío de un paquete de control es computado como un paquete nuevo. La normalización implica la división de esta suma total por el número de paquetes enviados.

Para caracterizar cada uno de los escenarios simulados se utiliza la duración media del enlace, tal y como se recomienda en [12]. La duración de un enlace se calcula midiendo el tiempo en el que dos nodos permanecen dentro de su mismo radio de cobertura, es decir, cuando la comunicación entre ambos puede establecerse directamente. Para esta medida se emplean todos los enlaces de la red, independientemente de que no vayan a ser empleados para la transmisión de los datos.

El propósito de las simulaciones es analizar la influencia que ejerce el intervalo de envío de MRA. Para ello se fijó un intervalo de RA equivalente a 2 segundos, variando en todas las simulaciones el intervalo de MRA en 2, 4 y 6 segundos.

Con el objetivo de presentar resultados más fáciles de entender se han interpolado los resultados, obteniéndose las líneas continuas.

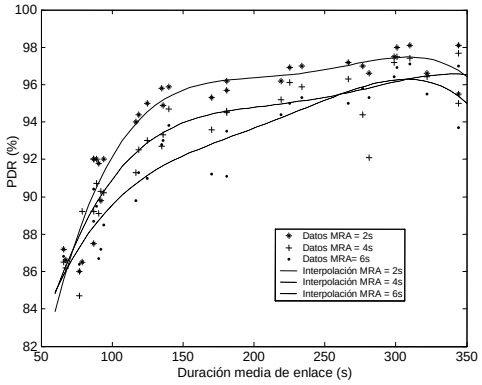


Figura 3. Packet Delivery Ratio (PDR).

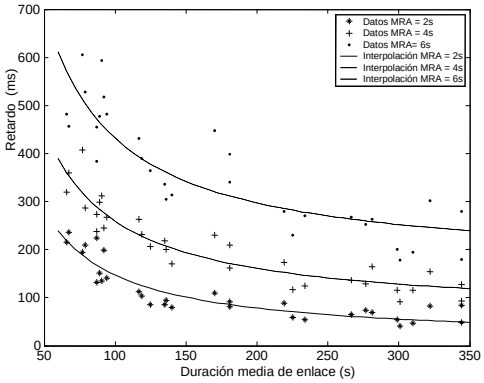


Figura 4. Retardo medio resultante.

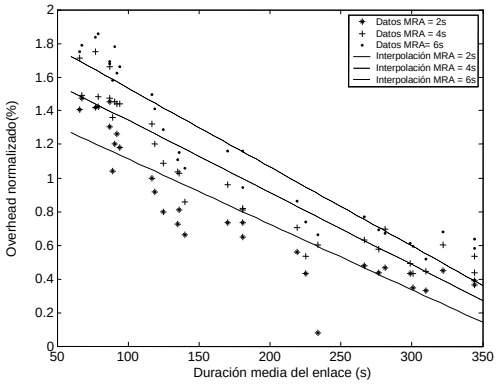


Figura 5. *Overhead* normalizado.

En líneas generales se aprecia que para escenarios con una mayor estabilidad, es decir, donde la duración del enlace es mayor, se obtienen mejores prestaciones. Bajo estas condiciones, los terminales van a mantener rutas estables durante más tiempo y con ello se reducirá el número de procedimientos de recuperación tras la ruptura de las rutas. Así pues, la Fig. 3 muestra que hay menos pérdidas de paquetes

ya que existe una menor probabilidad de colisiones entre paquetes al estar la red menos cargada con paquetes de control, como refleja la Fig. 5. Por otro lado, en el retardo se aprecia que posee una tendencia contraria a la duración del enlace. Ciertamente cuando las rutas son estables el paquete puede dirigirse al destino sin la necesidad de ser almacenado en ninguna cola ni de hacer ningún tipo de descubrimiento de ruta.

En aquellos escenarios donde la movilidad de los nodos implique un mayor número de rupturas de enlace va a ser necesario proceder al descubrimiento de las nuevas rutas. Como el tráfico que se ha analizado va dirigido hacia un nodo de Internet, se generarán paquetes *broadcast* de MRS para conocer cómo se puede alcanzar el *gateway*. Mientras se recibe la respuesta del *gateway* a través de MRA *unicast,* los terminales que se encuentren en esta situación deberán almacenar los datos en una cola interna de tamaño limitado a 64 paquetes. En el caso de que la capacidad de este almacenamiento fuese insuficiente o bien que el tiempo de almacenamiento resultase excesivo, el nodo empezará a descartar los paquetes y no procederá a su posterior envío. Estos son los motivos fundamentales que justifican que para duraciones de enlace bajas exista una mayor carga en la red, el retardo aumente y el PDR disminuya.

Para comprender las diferencias que ocasiona la elección del intervalo de envío de MRA es conveniente centrarse en la conmutación de los *gateways*. Tal y como se ha explicado anteriormente, en este mecanismo los *gateways* no están condicionados a permanecer dentro del área de cobertura del *router* de acceso. Al ser el envío de RA periódico, el *gateway* por defecto es capaz de detectar que se ha alejado de dicha zona. En ese instante dejará de realizar las funciones que tenía asociadas. Sin embargo, los *gateways* candidatos no iniciarán el procedimiento de autoconfiguración hasta que no haya trascurrido el periodo previsto para la recepción del mensaje de MRA. Por lo tanto, existen múltiples periodos a lo largo de la simulación en los que no existe ningún *gateway* configurado. Durante estos periodos las fuentes van a detectar que las rutas hacia Internet no están activas, almacenarán los datos en las colas internas y preguntarán por el *gateway* a través del envío de MRS. Es evidente que la duración de los periodos en los que no existe *gateway* es proporcional al intervalo elegido para la emisión de MRAs.

La Fig. 4 muestra la influencia que tienen esos periodos en el retardo. Los retardos medios medidos para intervalos de MRA equivalentes a 6 segundos son mayores que los que se obtienen para un intervalo de 4 segundos, que a su vez sobrepasa al retardo asociado a una elección de MRA igual a 2 segundos.

Esta dependencia va a influir de manera distinta en el PDR. La Fig. 3 muestra claramente que en los escenarios donde la duración del enlace es mayor, cuando los intervalos son menores se obtienen mayor tasa de paquetes recibidos. Sin embargo para entornos de alta movilidad fundamentalmente va a predominar la pérdida de paquetes bien sea por el llenado de la cola interna de los nodos o por las colisiones que acaecen en un medio inalámbrico donde se han incrementado los paquetes de control debido, precisamente, a esa inestabilidad en las rutas.

La Fig. 5 puede parecer en un principio contradictoria pues eligiendo un intervalo de MRA mayor cabría esperar un menor número de paquetes de control. Sin embargo no hay que omitir el hecho de que las fuentes están continuamente emitiendo. Por lo tanto demandan rutas actualizadas, procediendo al descubrimiento de las mismas al detectar que son obsoletas. Con el envío de un MRA *broadcast,* que es el que se realiza periódicamente, se actualiza la ruta hacia Internet de todos los nodos de la MANET. Sin embargo las peticiones *broadcast* que envía cada una de las fuentes reciben una respuesta *unicast* que sólo informará a la fuente origen y a aquellos nodos intermedios por los que se propague el MRA *unicast.* Por lo tanto, la carga ocasionada en la red es mayor que con el envío de MRA *broadcast* periódicos. Este comportamiento se apreciará en mayor medida cuando más fuentes carezcan de rutas actualizadas hacia Internet así como cuando dicha carencia sea más prolongada.

Analizadas estas tres gráficas, se puede concluir que la elección de un intervalo de envío de MRA equivalente a dos segundos proporciona el mayor número de ventajas en el rango de movilidad estudiado.

5 Conclusiones

Se ha mostrado el comportamiento del mecanismo que posibilita el acceso a Internet en una red ad hoc a través de múltiples *gateways* móviles.

Mediante el uso de simulaciones se ha concluído que el mayor inconveniente de este mecanismo reside en el hecho de que la autoconfiguración dinámica de *gateway* implica la existencia de periodos de tiempo en los que no existe ningún gateway por defecto configurado, y por lo tanto, los nodos son incapaces de enviar paquetes a Internet. Cabría esperar, pues, que el desarrollo de nuevas estrategias que reduzcan la duración de estos intervalos podrían repercutir en una mejora de las prestaciones.

Agradecimientos

La autora principal agradece la invitación del Instituto Tecnológico Avanzado de Samsung en Seúl.

También agradece los inestimables comentarios de Shubhranshu Singh y de Jae Hoon Kim.

Este trabajo ha sido parcialmente costeado por el proyecto de financiación pública Nº TEL2003-07953-C02-01.

Referencias

[1] T. Narten, E. Nordmark, W. Simpson. "Neighbor Discovery for IP version 6". RFC 2461. Diciembre 1998.

[2] C.E. Perkins, T. Marinen, R. Wakikawa, E.M. Beilding-Royer, Y. Sun. "IP address autoconfiguration for ad hoc networks". IETF Internet Draft, noviembre 2001. Trabajo en progreso.

[3] C. Huitema, "IPv6: the new Internet Protocol", Prentice-Hall, 1998. ISBN:0-13-850505-5.

[4] R. Wakikawa, J. Marinen, C. Perkins, A. Nilsson, A.J. Tuominen. "Global Connectivity for IPv6 Mobile Ad hoc Networks". IETF Internet Draft, oct. 2003. Trabajo en progreso.

[5] C. Jelger, T. Noel, A.Frey. "Gateway and address autoconfiguration for IPv6 adhoc networks". IETF Internet Draft, oct. 2003. Trabajo en progreso.

[6] S. Singh, J. H. Kim, Y.G. Choi, K.L. Kang, Y.S. Roh. "Mobile multi-gateway support for IPv6 mobile ad hoc networks". IETF Internet Draft, junio 2004. Trabajo en progreso.

[7] S. Deery, R. Hinden. "RFC 2460 – Internet Protocol, Version 6 (IPv6) Specification ". IETF Internet RFC, noviembre 1998.

[8] Alex Ali Hamidian. "A study of Internet Connectivity for Mobile Ad Hoc Networks in NS 2". Master´s thesis, Lund Institute of Technology, enero 2003.

[9] K. Fall, K. Varadhan "Ns Notes and Documentation", The VINT Project. UCBerkey, LBN, 2003.

[10] J. Yoon, M. Liu, B. Noble. "Random waypoint considered harmful". Proceedings of Infocom'03, pp. 1312-1321. San Francisco, abril 2003.

[11] C. Perkins, E. Beilding-Royer, S.Das "RFC 3561 - Ad hoc On-Demand Distance Vector (AODV) Routing". IETF Internet RFC, julio 2003.

[12] J. Boleng, W. Navidi, T. Camp, "Metrics to enable Adaptative Protocols for Mobile Ad Hoc Networks". Proceedings of ICWN'02, pp. 293-298. Junio 2002

[13] E. Casilari, A. Triviño Cabrera, "A practical study of the Random Waypoint mobility model in simulations of ad hoc networks", aceptado para The 19th International Teletraffic Congress (ITC19), Beijing (China), 29-Agosto-2 Septiembre, 2005

.

Evaluación de los sistemas inalámbricos 802.11b en entornos de alta movilidad para su implantación en una red de Infostations

Juan Ramón Cayón Alcalde y Eduardo Magaña Lizarrondo
Departamento de Automática y Computación
Universidad Pública de Navarra
C/ Campus Arrosadia, 31006 Pamplona
Teléfono: 948 169853 Fax: 948 168924
E-mail: eduardo.magana@unavarra.es

Abstract*. Nowadays, users demand higher bandwidth Internet access with the ability to connect anywhere (or, at least, many-where) at anytime, with no restrictions imposed by mobility or the kind of information (voice or data) to exchange. In this context, we've tried to explore the possibilities of a largely characterized and cheap wireless solution, 802.11b, in a high mobility environment. The performance of 802.11b wireless LANs is well known for indoor and static environments, while their behavior in outdoor and mobile environments has not been explored but from a theoretical point of view. Our work tests an empirical approach to 802.11b real possibilities on a high speed mobile environment. The goal was to check if a mobile station, moving at vehicular speed, was able to communicate with an AP at the roadside. Our results show that it is possible and achieving interesting throughput levels.*

1 Introducción

Los sistemas 802.11 [1] no fueron diseñados para trabajar en entornos móviles, sin embargo la evolución de la demanda del mercado de las telecomunicaciones ha ido obligando a la búsqueda de vías que permitan que servicios que hasta ahora sólo estaban disponibles en entornos estáticos, sean accesibles al mundo de las comunicaciones móviles. El objetivo es intentar que las redes inalámbricas cubran cada vez superficies más amplias, dando servicio tanto a usuarios en movimiento como estáticos [2]. La forma de hacerlo es integrando distintos protocolos inalámbricos en un mismo sistema, de modo que se satisfagan todos los tipos de demanda de servicio, de una manera totalmente transparente al usuario. Es en este escenario donde tecnologías ya existentes y de bajo coste, como es el caso de 802.11b, pueden reclamar su espacio si su rendimiento nos ofrece una QoS aceptable.

El comportamiento y rendimiento de 802.11b en entornos estáticos ha quedado totalmente caracterizado a partir de la gran cantidad de estudios existentes, tanto teóricos como prácticos [3]. También se ha estudiado el comportamiento en entornos de baja movilidad; en concreto el proceso de handoff entre distintos APs [4], cuando trabajamos en modo infraestructura, y el desarrollo de redes móviles Adhoc [5]. Para el caso de entornos de alta movilidad se han realizado estudios mediante emuladores [6], sin embargo no tenemos referencias de que exista ningún estudio empírico al respecto. Las únicas referencias existentes pertenecen a iniciativas privadas o con capital privado [7][8][9], que implementan servicios y realizan experiencias piloto, en muchas ocasiones utilizando soluciones propietarias, y sobre los que no se publican resultados. La ausencia de datos de medidas de campo lleva a plantearse de nuevo una pregunta que otros se habían hecho con anterioridad a nivel teórico: cuáles son los límites “reales” de movilidad de la tecnología WLAN.

2 El Concepto de *Infostation*

El modelo de red celular, ampliamente extendido gracias a la telefonía móvil, presenta una estructura que permite un acceso ubicuo a los servicios ofrecidos por la misma. Dichos sistemas están diseñados de modo que cualquier usuario dentro de una célula, independientemente de su distancia a la estación base de la misma, obtenga una mínima calidad en el servicio. Esto, unido a la continuidad de la estructura celular, da como resultado una cobertura en cualquier momento y en cualquier lugar.

Frente al modelo de red celular, con cobertura ubicua y células adyacentes pero ineficiente desde el punto de vista del *throughput/coste*, surge el modelo de red de *infostations*. Esta arquitectura de red, ya enunciada en 1996 [10], se caracteriza por una estructura discontinua de celdas aisladas y separadas entre sí, dentro de las cuales existe un gran ancho de banda que permite la descarga de grandes ráfagas de datos durante el tiempo que el móvil las atraviesa. Un diseño de estas características resulta eficiente porque los nodos se comunican sólo cuando están próximos y tienen unas buenas condiciones de canal. Si bien es cierto que el tamaño de las islas de cobertura es pequeño, favorece el empleo de bajos niveles de potencia en la estación móvil y de bandas de frecuencia distintas a las de la telefonía móvil.

Al contrario de lo que ocurre con las comunicaciones de voz, la transmisión de datos de determinadas aplicaciones es mucho más tolerante frente a retardos

y no requiere una conexión continuada ni una velocidad de transmisión específica, por lo que encajaría mejor en este concepto de sistema inalámbrico eficiente, en el que pasamos de una cobertura en todo instante y en cualquier lugar a una cobertura en algunos instantes y en determinados lugares.

Dentro de los distintos escenarios en los que podríamos aplicar esta estructura de red, resultan especialmente interesantes aquellos en los que existe movilidad por parte del usuario y particularmente movilidad a altas velocidades. En estos casos, dado el poco tiempo que pasa el móvil dentro de la zona de cobertura, se podrá dar la circunstancia de que la descarga no se complete por lo que deberá continuar cuando el usuario atraviese otra celda (Figura 1).

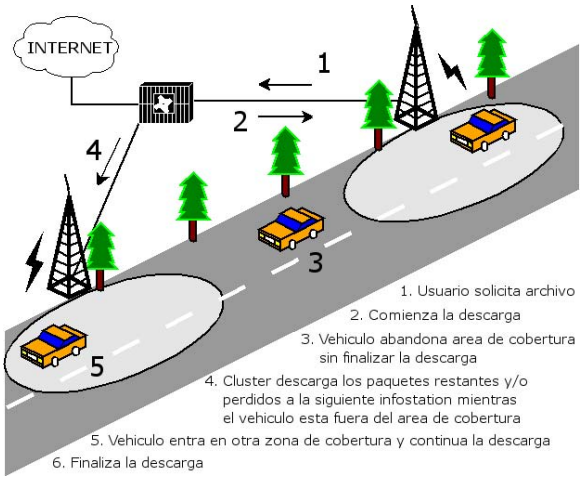


Figura 1: Cluster de infostations para clientes móviles de alta velocidad

Evidentemente, este tipo de estructura de red plantea nuevos retos y necesidades, como pueden ser la predicción de movilidad de la estación, o incluso la predicción de la información que pueda solicitar un usuario, necesidades que requieren la utilización de protocolos más complicados a nivel de gestión de red y de datos, para los que ya se han propuesto algunas soluciones concretas [11].

Las posibles aplicaciones de estos sistemas son múltiples: descarga de archivos multimedia (audio o vídeo), descarga de correo electrónico, de información sobre el estado de la red viaria, mapas (formato plano, ortofotos, renderizaciones 3D) de la zona por la que se esté viajando, etc. Existen proyectos de aplicaciones concretas, en su mayoría basados en el estándar DSRC (Dedicated Short Range Communications) [12], como pueden ser el DriveBy InfoFuelling de la Daimler Chrysler [13] o Infostations for Emergency Applications del WINLAB de la Rutgers University [14]. Escenarios como una autopista o un tren, son especialmente interesantes, ya que poseen un trazado concreto y, por tanto, un desplazamiento fácilmente predecible. El aprovechamiento de esta circunstancia permite una mayor eficiencia a la hora de servir la información.

3 Diseño del experimento

3.1 Equipo

Para la construcción de la estación móvil empleamos un Volkswagen Passat, en cuyo interior montamos un ordenador portátil DELL Inspiron 8100 (procesador Intel Pentium III Mobile a 1GHz y 256Mb de RAM) con una tarjeta NL-2511CD PLUS EXT configurada en modo Managed (cliente). La tarjeta, a su vez, estaba conectada a una antena externa montada en el techo del vehículo.

El nodo estático, lo construimos colocando sobre una mesa con ruedas (en la que iba montada la antena externa sobre un bastidor) un ordenador portátil DELL Latitude (procesador Pentium III 500MHz y 256Mb de RAM), con una tarjeta NL-2511CD PLUS EXT configurada en modo Master (hostAP). De este montaje resulta un punto de acceso ligero y de fácil transporte que podemos ver en la Figura 2.

Uno de los mayores problemas a superar está en la implementación tecnológica. Todos los estudios empíricos realizados muestran cómo el comportamiento de los sistemas 802.11 es altamente dependiente del hardware empleado. Los equipos presentan eficiencias muy diferentes tanto a nivel individual [3], como a la hora de interactuar para implementar una red [4].

Figura 2: Punto de acceso empleado durante las medidas de campo.

Aunque el empleo de equipos de un mismo fabricante no tiene porqué asegurar un mejor rendimiento del sistema [4], utilizamos tarjetas idénticas para implementar la estación móvil y el punto de acceso por varios motivos: (1) *su alta potencia de salida*, pues tienen una potencia de emisión de 20dB frente a los 12-15dB que suelen tener las tarjetas normales; (2) *el conector de antena externa*, que permite aumentar la ganancia; (3) *el modo Access Point*, que permite configurar la tarjeta para funcionar en modo punto de acceso.

El SO empleado en ambos equipos fue Red Hat Linux 9 con kernel 2.4.20-13.9.HOSTAP [15]. Para poder hacer funcionar las tarjetas en modo AP en Linux es necesario el driver hostAP [16], por lo que el kernel que empleamos es una versión ya compilada que contiene dichos drivers instalados. Como herramienta de generación de tráfico no se ha podido utilizar herramientas como *iperf* porque necesitan establece una conexión de control paralela que los hace no operativos para nuestro caso de estudio en el que fuera del área de cobertura no se puede establecer tal conexión de control. Por tanto realizamos nuestro propio generador de tráfico UDP. Dicha aplicación nos permite controlar la tasa de transmisión, el tamaño de paquete y el puerto y máquina destino, y nos muestra por pantalla el identificador y *timestamp* de los paquetes que envía/recibe y la tasa media de transmisión cada segundo.

3.2 Escenario de medida

Se llevaron a cabo una serie de trabajos previos en el laboratorio, de cara a tratar de determinar si nuestro experimento era viable, de qué modo podíamos ejecutarlo e identificar cuáles iban a ser los principales obstáculos con que nos íbamos a encontrar. Un experimento de estas características presenta distintos problemas de tipo técnico, que dificultan enormemente la realización de medidas repetitivas. Por ello, se ha tratado de sistematizar y acotar al máximo las condiciones de trabajo, de modo que podamos dotar de un mínimo de validez a los resultados obtenidos.

Nuestro objetivo principal es comprobar si una estación móvil, desplazándose a una velocidad superior a los 50km/h era capaz de comunicarse con un AP situado junto a la calzada. Simplificando este planteamiento al máximo, se trataba de que una estación generase tráfico, la otra lo captase y, para ver la bondad de la comunicación, analizar cuánto del tráfico generado había sido capturado.

Para las medidas de campo es vital la elección del lugar donde se realizarían. Es necesario poder alcanzar una cierta velocidad de desplazamiento significativa, pero también es importante incluir en la medida la degradación multicamino, que supone la presencia de edificios y otros obstáculos. Se eligió el tramo de carretera que muestra la Figura 3 porque reúne todas las condiciones anteriores: se pueden alcanzar velocidades de hasta 80 km/h y el tramo de cobertura presenta un edificio de unos 8 metros de altura a un lado de la carretera y un talud de similares dimensiones al otro, con farolas, árboles y arbustos, tanto a ambos lados de la vía como en la mediana. Por tratarse de una zona con tráfico medio, la constante presencia de coches también supone la adición de nuevos obstáculos, que complican las condiciones de conexión. Además, por tratarse de un tramo situado entre dos rotondas, posibilita la configuración de un "circuito" que permite circular al vehículo de manera ininterrumpida y a una velocidad constante, al menos dentro de las zonas de cobertura.

En cada captura el vehículo da dos vueltas completas al circuito, pasando 4 veces por la zona de cobertura. Lógicamente, cuando circula en sentido sur-norte, al estar un poco más alejado del AP y haber más obstáculos interpuestos sucede que la conexión se ve ligeramente degradada con respecto a cuando circula en sentido norte-sur. Al no poder realizar medidas en una zona de autopista nuestro rango de velocidades se ve limitado. Resultados obtenidos en estudios teóricos previos [6], indicaban que, con una SNR de 20dB el throughput se estabiliza en torno a los 20km/h y con una SNR de 15dB en torno a los 40 km/h, cayendo de los 3Mbps de la primera a 1Mbps en la segunda. Así pues seleccionamos velocidades de 60 km/h y 80km/h, lo suficientemente altas como para estudiar el comportamiento de un sistema que se supone diseñado para entornos estáticos.

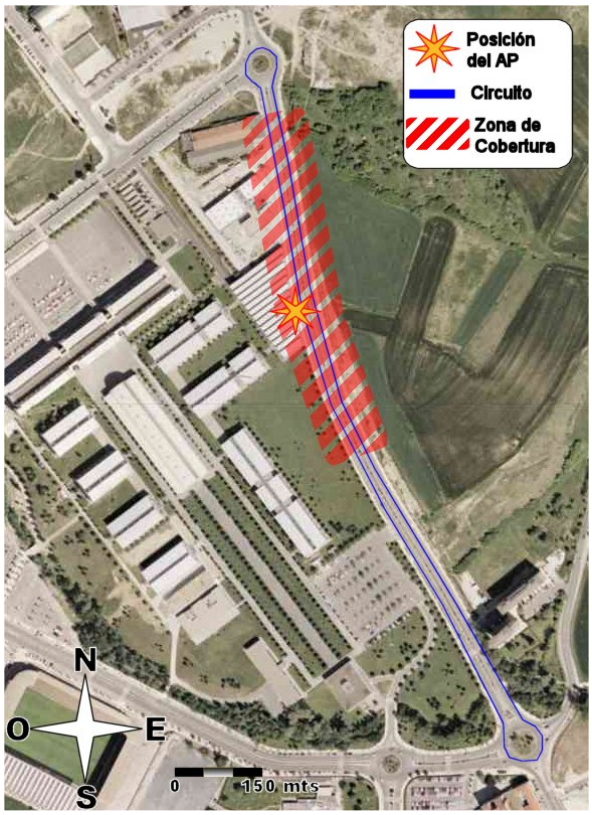


Figura 3: Vista aérea de la zona de medidas.

Para la configuración de los equipos creamos nuestra propia red, a la que identificamos bajo el nombre de *experimento*, establecemos el canal 6 como frecuencia de trabajo (2,437GHz) y transmitimos a la máxima potencia (20dB) dejando uno como estación móvil o cliente (*managed*) y el otro como estación base o AP (*master*). Para el resto de parámetros dejamos la configuración por defecto que traen las tarjetas, con lo que la modulación es seleccionada automáticamente por las mismas.

3.3 Medidas

Durante todo el experimento la estación móvil ha funcionado como fuente de tráfico y la estación base, o AP, como receptor del mismo. Hemos utilizado nuestros propios programas fuente y sumidero, basados en tráfico UDP, para generar el flujo de paquetes entre ambos equipos.

La fuente, genera un envío de paquetes en el que controlamos tanto el tamaño de los mismos como la velocidad a la que se transmiten. Esto nos permitirá, a posteriori, realizar comparativas del comportamiento del sistema a distintas tasas de transmisión y para distintos tamaños de paquete. Todos los paquetes contienen un índice identificador y el timestamp de cuando han sido generados, los cuales quedan almacenados en un archivo de datos. Además, la fuente realiza un cálculo sobre la tasa real a la que está enviando los paquetes, en Mbps. Es importante hacer notar que la tasa de transmisión es la velocidad a la que la fuente genera los paquetes y los entrega al interfaz para su envío siendo el interfaz el que, según las condiciones del canal, decide si enviar o no los paquetes y qué velocidad de modulación emplear para ello (1, 2, 5.5 o 11 Mbps para 802.11b). La selección de la modulación se ha dejado en manos de la propia tarjeta. Los resultados de simulaciones previas [6] muestran que, en general, se obtienen mayores niveles de throughput y menores pérdidas de paquetes en el modo automático que en el resto de los modos de manera aislada.

El sumidero se queda escuchando en un puerto concreto, recoge los paquetes, extrae el identificador de paquete, almacenando en el archivo de datos el identificador y el timestamp de ese instante. La decisión de hacer las medidas en un entorno con topología en modo infraestructura (con la presencia de un AP) implica que para poder comunicarse, la estación móvil ha tenido que, previamente, detectar la presencia del AP y asociarse al mismo. De este modo, los retardos producidos en el proceso de asociación ya están incluidos en la medida.

Para la realización del experimento hemos considerado las tasas de transmisión comprendidas entre 500 Kbps y 4Mbps, con saltos de 500Kbps; los tamaños de paquete empleados han sido: 64, 500, 1000 y 1500 bytes.

3.4 Zonas de *cobertura aceptable*: mesetas

Para el procesado de las medidas hemos empleado una serie de scripts, a partir de los cuales obtenemos, para cada segundo, el número de paquetes recibidos y perdidos y el throughput estimado y real. Conociendo las condiciones de tasa de transmisión, tamaño de paquete, etc., de la captura que procesemos, podemos obtener la información recibida y perdida por segundo, en distintas unidades: en %, en Kbps, etc.

Inicialmente procesamos el número de paquetes recibidos por segundo, cruzando los datos obtenidos en los archivos de captura de origen y destino, y tomando como referencia temporal el *timestamp* del primero: en el archivo de la fuente vemos qué paquetes han sido generados dentro de un determinado segundo y, con esta información, vamos al del destino y comprobamos cuántos de éstos han sido recibidos. La Figura 4 muestra el resultado.

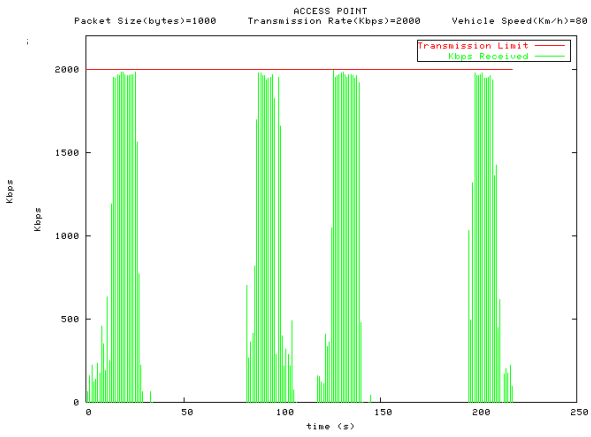


Figura 4: Información "entregada" (Kbps) frente al tiempo, durante todo el intervalo de medida.

Así pues, el efecto final es que no estamos midiendo tanto la información recibida por segundo sino, más bien, cuanta de la información generada en un determinado segundo llega a su destino.

Por otro lado, no es lógico considerar que tenemos cobertura, por el simple hecho de que recibamos uno o dos paquetes, en un segundo determinado. Es necesario discriminar parte de la información capturada, y aquí es donde entra nuestro concepto de *cobertura aceptable*. Realmente no es un concepto sino el conjunto de restricciones que hemos impuesto, a la hora de procesar los datos, para definir unos intervalos de cobertura más o menos homogénea. Dichas condiciones son:

- Sólo consideraremos que hay cobertura en aquellos segundos en los que recibamos, al menos, el 90% de la información transmitida.
- El intervalo de cobertura ha de tener una duración mínima de 5 segundos
- Dentro del intervalo de cobertura puede haber vanos (segundos dentro de los cuales no recibimos el 90% de la información) pero no pueden tener más de 2 segundos de duración.

Una vez incluidas las restricciones, obtenemos el resultado mostrado en la Figura 5. Pueden observarse cuatro intervalos de cobertura, claramente identificables, que se corresponden con los cuatro pasos del vehículo por el entorno de la estación base, y que hemos dado en llamar *mesetas*. En realidad, lo que estamos mostrando es el throughput o, más

concretamente, durante cuantos segundos recibimos, al menos, el 90% de la información transmitida

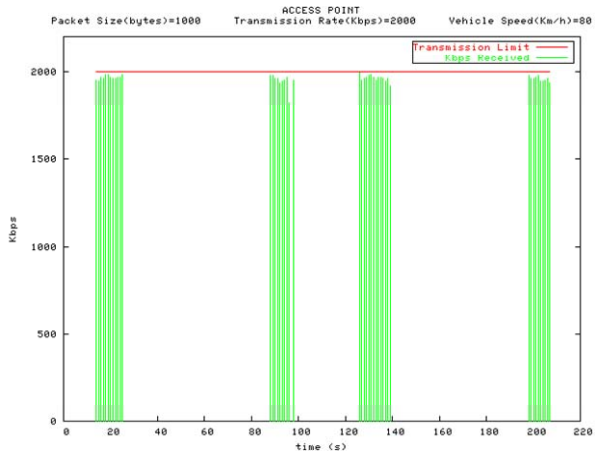


Figura 5: Información "entregada" (Kbps) frente al tiempo, tras aplicar a los datos de la las restricciones de cobertura aceptable

4 Resultados y análisis

Quizás el número de capturas realizadas no sea lo suficientemente amplio como para lanzar afirmaciones categóricas sobre el rendimiento del protocolo 802.11b en entornos de alta movilidad. Sin embargo, los resultados que a continuación presentamos nos parecen muy interesantes y muestran un comportamiento que nos permite hacer conclusiones de interés.

A la hora de estudiar los resultados de las distintas medidas, hemos de considerar las tres variables con las que hemos jugado, esto es: tasa de transmisión, tamaño de paquete y velocidad del vehículo o estación móvil. Para facilitar el análisis comparativo del comportamiento del sistema hemos aislado los estudios para cada una de las velocidades de vehículo (60 y 80 km/h) viendo cómo variaban el nivel de información recibida y perdida (en tanto por ciento) y la duración de las *mesetas* o zonas de cobertura (en segundos), tanto en función de la tasa de transmisión como en función del tamaño de paquete. Finalmente hemos contrastado los resultados de ambos estudios, para ver en qué modo ha afectado la velocidad de la estación móvil.

Las comparativas se han realizado para los cuatro tamaños de paquete empleados en las medidas; sin embargo, las capturas realizadas con paquetes de 64 bytes tan sólo ofrecen resultados satisfactorios dentro de los parámetros de *cobertura aceptable* por nosotros establecidos, a 60km/h y 500Kbps de tasa de transmisión. A mayores tasas el nivel de información recibida está siempre por debajo del 90% de la transmitida. Si nos fijamos en el método de acceso al medio de 802.11 veremos que, una vez establecida la comunicación entre estación móvil y AP, cada paquete de datos enviado se responde con un ACK desde el receptor. Con un tamaño de paquete tan pequeño, estamos enviando muchos paquetes por segundo, con lo que el número de tramas correspondientes al protocolo de señalización 802.11b también será muy alto reduciéndose el ancho de banda disponible para el envío de información. De hecho, si pensamos que 802.11 está basado en Ethernet y que en dicho protocolo una trama no puede tener menos de 64 bytes, nos encontramos que cuando realizamos transmisiones con un tamaño de paquete de 64 bytes, el 50% de la información que circula por el canal se corresponde con tramas ACK, con lo que estamos reduciendo el ancho de banda disponible para datos a la mitad como mínimo.

Por otro lado, a partir de los 3.5Mbps de tasa de transmisión, ocurre lo mismo con los paquetes de 500 bytes; sin embargo con los paquetes 1000 bytes y 1500 bytes todavía obtenemos datos, tanto a 3.5Mbps como a 4Mbps, aunque los intervalos de cobertura se reducen drásticamente.

4.1 Efecto sobre el tiempo de cobertura

Para obtener la medida del tiempo de cobertura realizamos un conteo de los segundos durante los cuales se cumplen las condiciones que hemos definido como *cobertura aceptable*. El resultado comparativo podemos verlo en la Figura 6 para 60Km/h y en la Figura 7 para 80Km/h.

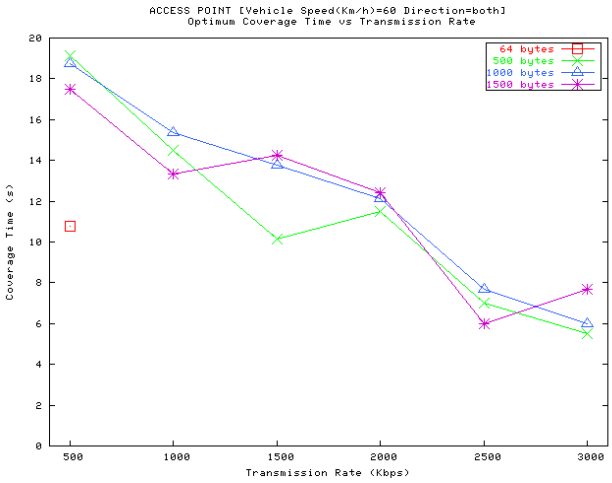


Figura 6: Intervalo de Cobertura (sg) vs Velocidad de Transmisión a 60 km/h

Cuanto mayor sea la tasa de transmisión mayor será la velocidad de modulación que he de emplear para el envío de la información. Sin embargo, la velocidad de modulación está regida por las condiciones del canal: cuanto mejores sean éstas, mayor velocidad de modulación podremos emplear. Por lo tanto, al aumentar la tasa de transmisión estamos exigiendo unas mejores condiciones de canal para poder establecer la comunicación. Esto se traduce en intervalos de cobertura más estrechos y, por extensión, en una mejor SNR promedio dentro de dicho intervalo, como veremos en el apartado siguiente. Por tanto, a menor tasa de transmisión tenemos mayor tiempo de cobertura.

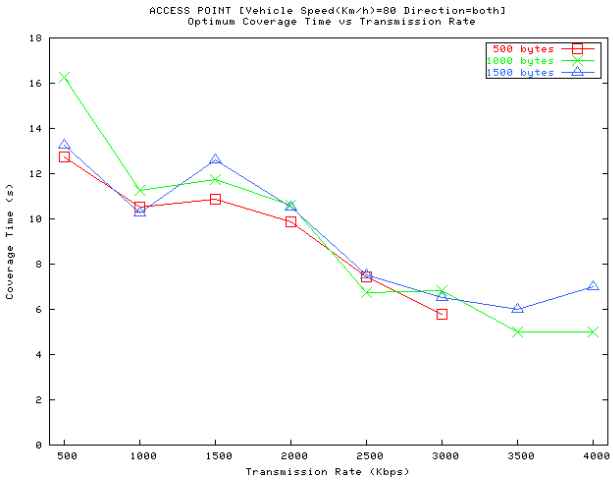


Figura 7: Intervalo de Cobertura (sg) vs Velocidad de Transmisión a 80 km/h

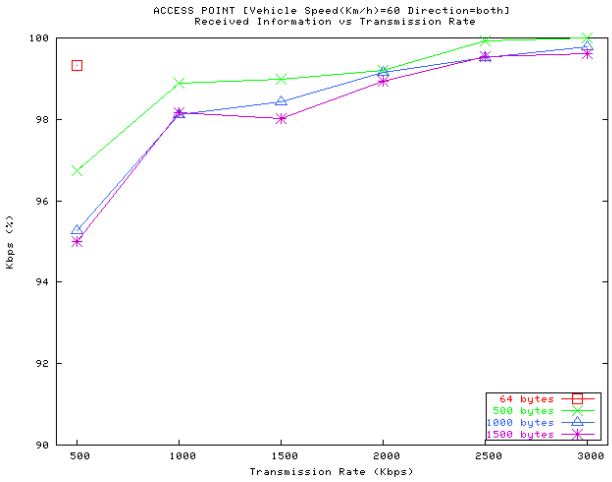


Figura 8: Información Recibida (%) vs Velocidad de Transmisión a 60 Km/h.

El efecto de la variación del tamaño de paquete no está tan claro; si bien parece que para tamaños de 500 y 1000 bytes el comportamiento es más o menos correlativo, presentando unos mejores resultados a 1000 bytes para ambas velocidades de vehículo, el comportamiento con paquetes de 1500 bytes es demasiado variable para poder emitir conclusiones definitivas.

Asimismo, tal y como cabía esperar, los tiempos de cobertura disminuyen al aumentar la velocidad del móvil, efecto lógico, ya que al aumentar la velocidad, tarda menos tiempo en atravesar la zona de cobertura.

4.2 Efecto sobre el throughput

Para estudiar el throughput hemos comparado el porcentaje de información recibida para las distintas tasas de transmisión y tamaños de paquete. En las figuras Figura 8 y Figura 9, puede apreciarse claramente cómo, a medida que aumenta la tasa de transmisión aumenta también el porcentaje de información recibida. Este hecho, que en principio puede parecer contradictorio, está perfectamente justificado.

Las gráficas muestran el porcentaje de información recibida, no en el total de la captura, sino dentro de la zona de cobertura y ésta es diferente en cada medida. Como ya hemos comentado en el apartado anterior, al aumentar la tasa de transmisión estamos exigiendo unas mejores condiciones de canal para poder establecer la comunicación. Esto se traduce en intervalos de cobertura más estrechos y, por extensión, en una mejor SNR promedio dentro de dicho intervalo. Los intervalos de cobertura quedan acotados en torno al AP y cuanto más cerca estemos del mismo mayor será la SNR que, por otro lado, disminuye muy rápidamente a medida que nos separamos del AP; así que cuanto más estrechos sean los intervalos mejores y más homogéneos serán los niveles de SNR.

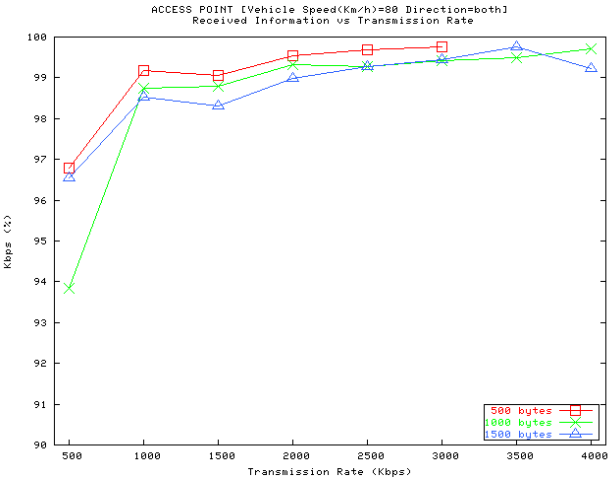
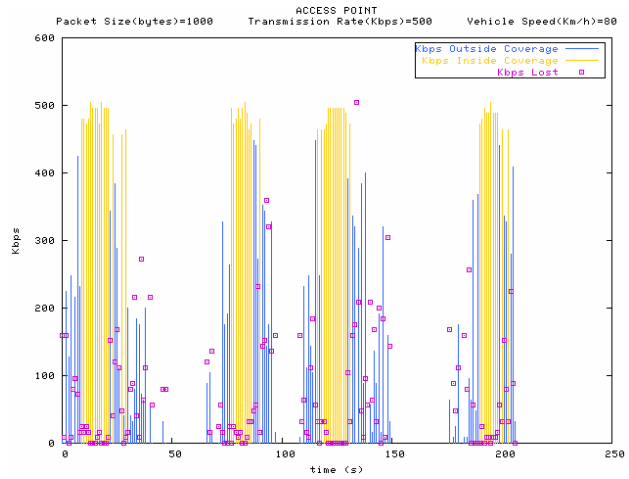


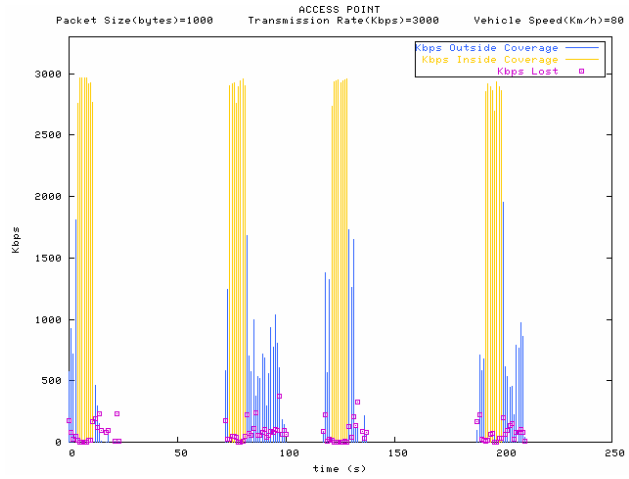
Figura 9: Información Recibida (%) vs Velocidad de Transmisión a 80 Km/h

En la Figura 10 se muestran dos capturas realizadas a 80 km/h con un tamaño de paquete de 1000 bytes y a tasas de transmisión de 500 y 3000 Kbps; se aprecia cómo se estrecha el intervalo de cobertura y cómo se acentúan las diferencias entre los paquetes recibidos dentro y fuera de la zona de cobertura aceptable. Asimismo podemos observar cómo los niveles de información perdida (número de paquetes perdidos) se mantienen más o menos similares en la zona central de las mesetas (las pérdidas se concentran en los bordes de las mesetas de la Figura 10a), lo que apoya el hecho de que porcentualmente la cantidad de información recibida aumente.

En cuanto al modo en que afecta la variación del tamaño de paquete, se aprecia con claridad cómo, para una misma tasa de transmisión, disminuye el porcentaje de información recibida a medida que aumentamos el tamaño de paquete. Es perfectamente lógico, pues cuanto mayor sea un paquete mayor es la probabilidad de pérdida al transmitirlo.



(a)



(b)

Figura 10: Estrechamiento de las zonas de cobertura con el aumento de la tasa de transmisión a) 500Kbps b) 3000Kbps

Finalmente, en la Tabla 1 vemos el promedio de los niveles de throughput que hemos alcanzado dentro de las mesetas, para cada velocidad del vehículo, tasa de transferencia y tamaño de paquete. Tanto a 60km/h como a 80km/h, y para todos los tamaños de paquete, coincide que obtenemos los mejores resultados para una tasa de transferencia de 2Mbps.

5 Conclusiones

En este trabajo hemos comprobado en qué modo se ve afectado el *throughput* por la tasa de transmisión y por el tamaño de paquete, constatando cómo el aumento de las pérdidas dentro del área de *cobertura aceptable* es directamente proporcional al tamaño de paquete e inversamente proporcional a la tasa de transmisión.

Asimismo hemos podido ver cómo los tiempos de cobertura disminuyen con el aumento de la tasa de transmisión, y aunque presentan un comportamiento más heterogéneo frente a las variaciones del tamaño de paquete, parece intuirse que el empleo de tamaños grandes ofrece mejores resultados.

En ambos casos, tanto el throughput como el tiempo de cobertura medidos, son coherentes con los resultados obtenidos en estudios previos mediante el uso de emuladores [6].

Hemos mostrado como los tiempos de cobertura son lo suficientemente largos como para permitir al cliente asociarse a un AP y transmitir una importante cantidad de paquetes. Por otro lado las limitaciones son patentes, pero siempre hay que tener en cuenta qué tipo de servicio estamos interesados en implementar y a qué coste. Esto nos lleva a afirmar que, en determinadas circunstancias la posibilidad de implementar servicios móviles bajo 802.11b podría ser viable. Si estamos buscando implementar servicios de bajo/medio peso en transmisión, como pueden ser servicios de información específica como por ejemplo sobre el estado de las carreteras (información vial, obras, desvíos alternativos, detección temprana de accidentes, ...), información turística local (mapas, rutas, hoteles, restaurantes, ...), etc., quizás 802.11b podría cubrir las expectativas necesarias. Incluso si consideramos aplicaciones de mayor peso, como la descarga de archivos multimedia, y teniendo en mente la estructura de una

60 Km/h

Vtx	TAMAÑOS DE PAQUETE 500			1000			1500		
	ambos	*norte-sur*	*sur-norte*	*ambos*	*norte-sur*	*sur-norte*	*ambos*	*norte-sur*	*sur-norte*
500	9,04	10,69	7,39	8,85	9,07	8,62	8,19	8,95	7,43
1000	14,09	15,03	13,14	14,72	15,63	13,82	12,87	14,89	11,26
1500	14,56	18,49	10,64	19,85	20,98	18,72	20,03	21,77	18,30
2000	**21,62**	**24,90**	18,35	**23,11**	**25,78**	20,44	**24,01**	**26,01**	21,36
2500	17,21	17,24	14,64	18,15	19,93	20,90	14,68	15,91	12,22
3000	16,28	17,73	14,84	17,72	17,71	17,73	22,40	16,82	22,72
3500	-	-	-	-	-	-	-	-	-
4000	-	-	-	-	-	-	-	-	-

80 Km/h

Vtx	TAMAÑOS DE PAQUETE 500			1000			1500		
	ambos	*norte-sur*	*sur-norte*	*ambos*	*norte-sur*	*sur-norte*	*ambos*	*norte-sur*	*sur-norte*
500	5,99	6,11	5,88	7,31	8,14	6,44	6,36	6,72	6,00
1000	10,08	10,56	9,70	11,03	11,27	10,67	10,05	10,78	9,22
1500	15,66	16,92	14,30	16,92	18,38	15,40	17,80	20,52	15,07
2000	**19,26**	**21,07**	17,37	**20,61**	**23,03**	18,05	**20,48**	**24,13**	16,98
2500	17,71	19,68	15,34	16,15	17,59	14,34	18,08	18,99	16,31
3000	16,59	16,85	15,72	19,88	19,51	20,58	19,11	18,38	20,58
3500	-	-	-	17,15	17,15	-	19,80	19,80	-
4000	-	-	-	18,95	18,95	-	**25,06**	25,06	-

Tabla 1. Mbits/meseta promedio obtenidos, en ambas velocidades de vehículo, para cada tasa de transmisión y tamaño de paquete

red de infostations (en la que, como ya hemos explicado, podemos predecir el movimiento de la estación), el uso de hardware 802.11b sería viable, siempre que, por encima, esa estructura de red tuviera los mecanismos apropiados (arquitectura y protocolos) que se encarguen del control y correcta distribución de los datos.

Los resultados expuestos en la Tabla 1 nos permiten ser optimistas sobre las posibilidades de implementación de 802.11b en aplicaciones reales. Si considerásemos, por ejemplo, un servicio de descarga de audio en formato MP3 (música para escuchar en el reproductor del vehículo o una descripción narrada del paraje natural-histórico-artístico que estemos atravesando), teniendo en cuenta que MP3 viene a tener un ratio aproximado de 1 minuto de reproducción por Mbyte de datos, podemos descargarnos en torno a unos 2,5 minutos de audio, lo cual nos permite además espaciar bastante el posicionamiento de las infostations. A 80km/h, bastaría tener una infostation cada 3,5 km para permitir una reproducción continua, sin retardos.

Como líneas futuras de trabajo será de interés revisar lo que ocurre en el otro sentido de la comunicación, esto es, haciendo que el AP funcione como fuente y la estación móvil como sumidero. Estudiar el funcionamiento de la red en modo Adhoc también podrá dar información interesante al eliminar el proceso de asociación con el AP. Finalmente, el efecto de las peculiaridades del canal sobre el protocolo de transporte TCP y sobre aplicaciones reales podrá permitir estudiar las necesidades de conversión de protocolos mediante proxys en una red de infostations sobre 802.11b.

Referencias

[1] IEEE. “IEEE 802.11 WLAN Working Group”

[2] Lee W. McKnight, J. Howison and S. Bradner. “Wireless Grids: Distributed Resource Sharing by Mobile, Nomadic and Fixed Devices”. IEEE Internet Computing, Jul./Aug. 2004.

[3] A. Vasan y A. Udaya Shankar. “An Empirical Characterization of Instantaneous Troughput in 802.11b WLANs”, Under Submission.

Christian Hoene, André Günther, Adam Wolisz. “Measuring the Impact of Slow User Motion on Packet Loss and Delay over IEEE 802.11b Wireless Links”. 28th Annual IEEE International Conference on Local Computer Networks, Bonn/Königswinter, October 2003.

[4] A. Mishra, M.Shin, W. Arbaugh. “An Empirical Analysis of the IEEE 802.11 MAC Layer Handoff Process”, ACM SIGCOMM Computer Communication Review, Vol. 33, Issue 2, April 2003.

[5] E. Welsh, P. Murphy, J. Patrick Frantz. “A Mobile Testbed for GPS-Based ITS/IVC and Ad Hoc Routing Experimentation”, International Symposium on Wireless Personal Multimedia Communications (WPMC), Vol. 2, Pg. 796-800, Honolulu, October 2002.

[6] C. Steger, P. Radosavljevic and J. Patrick Frantz. “Performance of IEEE 802.11b Wireless LAN in an Emulated Mobile Channel”, IEEE Vehicular Technology Conference (VTC), Jeju, Korea, April 2003.

[7] Keith Biesecker, “Broadband Wireless, Integrated Services And Their Application To Intelligent Transportation Systems”, Technical Report, Center for Telecommunications and Advanced Technology, McLean, Virginia. Junio 2000.

[8] Atheros. “Atheros Chipsets Used in Mercedes-Benz Future Technology Demonstration”, http://www.atheros.com/news/mercedesdemo.html

[9] Wixos, “Red de HotSpots en las líneas de Metro y Bus de París”, http://www.telcite.fr/nwifi3.htm

[10] R. H. Frenkiel and T.Imielinski, “Infostations: The Joy of ‘Many-Time, Many-Where’ Communications”. Tech. Rep. TR-119, Wireless Information Networks Laboratory (WINLAB), Rutgers State University of New Jersey, Apr. 1996.

[11] Ana Lúcia Iacono and Christopher Rose. “Infostations: New Perspectives on Wireless Data Networks”, Proceedings of NJIT Symposium on Next Generation Wireless Networks, Newark, NJ, May 2000.

[12] http://www.astm.org/, “E2213-03 Standard Specification for Telecommunications and Information Exchange Between Roadside and Vehicle Systems”, ASTM, Apr. 2003.

[13] C. Frank, W. Olfelder, D. Jiang, G. Matylis, P. Pepper. “Dynamic Software Architectures for a Sometimes Somewhere Telematics Concept”. Technische Universität Berlin y Daimler Chrysler RTNA, TR No.2003-11, Nov. 2003.

[14] Focus Projects. “Infostations for Emergency Applications”. http://www.winlab.rutgers.edu/pub/docs/focus

[15] Redhat RPMs con HostAP. http://www.cat.pdx.edu/~baera/redhat_hostap/

[16] Driver hostAP para tarjetas con chipset Prism de Intersil. http://hostap.epitest.fi/

Método Adaptable de Eliminación de Reenvíos Basado en Contador para Reducir la Sobrecarga de Datos en el Tráfico Multicast sobre Redes Ad Hoc

C. M. Yago Sánchez, P.M. Ruiz, A. F. Gómez Skarmeta
Departamento de Ingeniería de la Información y las Comunicaciones. Universidad de Murcia
Facultad de Informática. Campus Universitario de Espinardo.
30100– Murcia (España)
Teléfono: +34 986 36 46 40 Fax: +34 986 36 41 51
E-mail: {carmen, pedrom, skarmeta}@dif.um.es

Abstract*. In ad hoc networks, mesh based protocols have revealed as a way to provide reliability to routing multicast. But the alternate paths which make available this robustness also cause the delivering of a high number of redundant data messages. In this paper, we address the problem of reducing this data overhead while maintaining mesh based protocols performance. We propose a distributed mechanism named "mobility aware counter based mechanism" which is aware of network conditions to make every node decide if the packet it is going to forward is redundant or not. Results obtained through simulation demonstrate the reduction of data overhead while offering good protocol performance.*

1 Introducción

En la actualidad las tecnologías móviles e inalámbricas están experimentando un gran auge. En este contexto, las redes ad hoc o MANETs (*Mobile Ad hoc NETworks*) están despertando un gran interés. Estas redes están formadas por un grupo de nodos móviles que se comunican entre ellos sin que haya ninguna infraestructura fija. Estos nodos actúan simultáneamente como *host* y como *router* para lograr que cualquier mensaje llegue con éxito a su destino.

También existe un manifiesto interés por permitir a estas redes utilizar una comunicación multicast. De hecho, se han planteado diversos protocolos para encaminar este tipo de tráfico. Estos protocolos pueden dividirse básicamente en tres categorías [1]: protocolos sin estado como DDM [2], protocolos basados en árbol como MAODV [3] y protocolos basados en malla como ODMRP [4]. También existen propuestas híbridas como AMRoute [5].

Los protocolos sin estado se orientan a grupos multicast pequeños. Con grupos mayores se pueden utilizar los protocolos basados en árbol; no obstante, su rendimiento decae en entornos con alta movilidad. En este caso les superan los protocolos basados en malla, debido a que introducen más redundancia y caminos alternativos.

Dado que las redes ad hoc experimentan grandes y/o rápidos cambios, los protocolos basados en malla proporcionan una solución adecuada para encaminar de forma robusta tráfico multicast. Sin embargo, también introducen una considerable sobrecarga. Esta sobrecarga tiene dos causas: por un lado, la inestabilidad de la red obliga a inundarla periódicamente (*flooding*) con mensajes de control; por otro lado, encontramos la sobrecarga de datos. Esta sobrecarga se define como aquellos mensajes de datos transmitidos innecesariamente debido a la redundancia de la malla [6]. Es por tanto consecuencia del hecho de que la redundancia de la malla multicast proporciona robustez, pero a la misma vez hace que muchos paquetes de datos sean retransmitidos innecesariamente. Dado que el porcentaje de tráfico de datos es mayor que el porcentaje de tráfico de control, la sobrecarga de datos se convierte en la principal limitación de la eficiencia de los protocolos basados en malla, produciendo un consumo excesivo de ancho de banda y un aumento de la contienda en la capa de enlace.

Ruiz [6] demostró que el cálculo del árbol multicast con sobrecarga de datos mínima es un problema NP-completo. Por esta razón, consideramos que un buen método para aumentar la eficiencia manteniendo la robustez es el uso de un algoritmo aproximativo basado en la idea de limitar el número de mensajes de datos redundantes sin podar la malla multicast. Entre los algoritmos que presentan este comportamiento se encuentra el algoritmo basado en contador propuesto por Ni *et al*. en [7] para paliar el problema de la tormenta broadcast (*broadcast storm problem*). Basándonos en él, proponemos un mecanismo basado en contador que además se adapta a las condiciones de la red.

Este documento se organiza de la siguiente forma: en la próxima sección realizamos una breve descripción de las propuestas existentes en la literatura para reducir la sobrecarga tanto en mecanismos de inundación como en tráfico multicast. En la sección 3 proponemos nuestra

versión del algoritmo basado en contador: el mecanismo basado en contador adaptable a la movilidad, que es un método adaptable a las condiciones de la red para ser utilizado en encaminamiento multicast. En la sección 4 mostramos y evaluamos los resultados de las simulaciones que hemos llevado a cabo. Para finalizar, exponemos las conclusiones obtenidas.

2 Trabajo Relacionado

En la literatura relacionada con redes ad hoc encontramos un buen número de algoritmos y protocolos cuyo fin consiste limitar el número de mensajes vertidos en la red sin perder fiabilidad. Podemos dividir estas propuestas básicamente en dos categorías [8]: propuestas basadas en topología y propuestas basadas en heurística. La primera categoría utiliza información topológica para reducir el número de nodos que pueden retransmitir un mensaje. Por ejemplo, encontramos métodos que utilizan información de los vecinos como *Self-Pruning* [9] o *Multipoint Relay* MPR [10]. La segunda categoría sin embargo, no limita el número de nodos que pueden retransmitir, sino que cada nodo decide cada vez y en función de una heurística, si ha de reenviar o no el mensaje.

Dentro de la segunda categoría, Ni *et al* han definido en [7] tres algoritmos: basado en contador, basado en distancia y probabilístico. En el algoritmo basado en contador, un nodo no reenvía si ha escuchado mensajes duplicados más de un determinado número de veces. En el algoritmo basado en distancia, un nodo decide si retransmite o no el mensaje dependiendo de su distancia a otros nodos. En el algoritmo probabilístico, el nodo toma una decisión dependiendo de una función probabilística. Por ejemplo, el protocolo *Source Grouped Flooding* [11] para encaminamiento multicast utiliza una función probabilística para reducir el número de retransmisiones redundantes de mensajes de datos.

Consideramos que el algoritmo basado en contador es especialmente significativo debido a su rendimiento, su facilidad de uso y su bajo consumo de recursos. Se basa en el concepto de "cobertura adicional prevista" (*expected additional coverage*). Cada vez que un nodo retransmite un mensaje, éste cubre un área en la cual el mensaje puede ser escuchado por otro nodo. El concepto de "cobertura adicional prevista", se define como el área en la cual el mensaje se escucharía por primera vez si un nodo decide retransmitir el mensaje. Esta área se hace más pequeña cada vez que un nodo escucha repetido el mismo mensaje que tiene que reenviar. De hecho, después de haber escuchado el mismo mensaje tres veces o más, la "cobertura adicional prevista" está por debajo del 10% [7] del área de cobertura total del nodo.

Este concepto se muestra en el ejemplo de Fig. 1: los nodos A y B reenvían el mismo mensaje, el cual es escuchado por los nodos C y D. C recibe el mensaje dos veces (desde A y B) mientras que D sólo una vez (desde B). En ambos casos, la "cobertura adicional prevista" corresponde a la zona no sombreada de su área de cobertura, ya que cualquier nodo que estuviese situado en el área sombreada ya habría escuchado el mensaje proveniente desde A y/o desde B. Por tanto, si C o D retransmiten el mensaje, sólo los nodos que estuvieran en la parte no sombreada recibirían el mensaje por primera vez. La "cobertura adicional prevista" de C es más pequeña que la de D porque el área de cobertura de C ha sido casi totalmente cubierta por el área de cobertura de A y B, mientras que el área de cobertura de D ha sido sólo parcialmente cubierta por el área de cobertura de B.

El algoritmo basado en contador propone que un mensaje no debe ser retransmitido si ha sido escuchado M veces, siendo M el número de veces que hace que la "cobertura adicional prevista" sea lo suficientemente pequeña. Siguiendo con el ejemplo y considerando, para simplificar, $M=2$, D retransmitirá el mensaje ya que sólo lo ha escuchado una vez, mientras que C no lo hará ya que ha escuchado el mensaje dos veces.

Para hacer más eficiente la eliminación de mensajes redundantes, están apareciendo propuestas que adaptan estos algoritmos en función de alguna de las características de la red, como por ejemplo la densidad de vecinos. Siguiendo esta métrica, en [12] se definen los protocolos *Border Node Retransmission Based Probabilistic Broadcast Protocols* y en [13] se propone un algoritmo adaptable basado en contador para mecanismos de inundación.

Si se quiere que los protocolos adaptables presenten mejor rendimiento que los protocolos no adaptables, es necesaria una métrica representativa. Boleng, Navidi y Camp han definido en [14] los requisitos que debe cumplir una métrica de movilidad para servir de base a un protocolo adaptable. Según [14] una métrica ha de ser computable en un entorno real y distribuido, independiente de un protocolo específico y buena indicadora del rendimiento del protocolo.

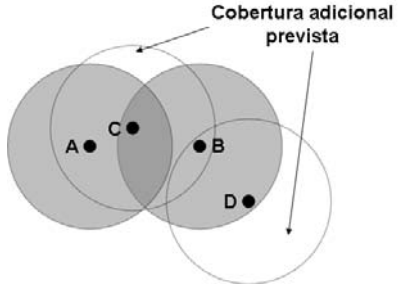


Figura 1: Ejemplo de "cobertura adicional prevista"

Basados en todo lo expuesto, proponemos aquí un método para reducir la sobrecarga de datos en protocolos multicast basados en malla: el mecanismo basado en contador adaptable a la movilidad. Consiste en una nueva versión mejorada del algoritmo basado en contador. Esta nueva versión se adapta a las condiciones de la red utilizando una nueva métrica, el intervalo modal de la duración del enlace. Esta nueva métrica se basa en la estabilidad de los enlaces entre los nodos, lo que permite al nodo conocer las condiciones locales de su entorno y reaccionar ante ellas.

3 Mecanismo Basado en Contador Adaptable a la Movilidad

Como se ha visto en la sección anterior, los mecanismos basados en contador ([7], [13]) resultan eficaces para reducir la redundancia cuando se utilizan con *flooding*. Nosotros empleamos una variante de este algoritmo para la reducción de la sobrecarga de datos en protocolos multicast basados en malla.

El funcionamiento del algoritmo basado en contador es el siguiente: cada nodo perteneciente a la malla dispone de un contador *c*. Cuando ha de reenviar un mensaje establece un temporizador y mientras que éste no finaliza, incrementa el contador *c* cada vez que escucha el mismo mensaje que él tiene que retransmitir. Cuando el temporizador expira, si el contador *c* ha alcanzado un valor umbral *C*, el mensaje de datos no se retransmite, es descartado. La justificación de este comportamiento, tal y como se ha comentado en la sección anterior, radica en el concepto de "cobertura adicional prevista".

En el algoritmo basado en contador, el valor umbral *C* representa el número estimado de mensajes que han de ser escuchados para que la "cobertura adicional prevista" sea tan pequeña que no merezca la pena introducir más redundancia. Es necesario elegir cuidadosamente este umbral: debe encontrarse un compromiso entre el rendimiento del protocolo y la reducción de la sobrecarga. Un mayor umbral *C* puede hacer que el número de mensajes correctamente entregados sea mayor, pero a costa de aumentar significativamente la sobrecarga.

Además, cuando los enlaces entre los nodos de la red son inestables, el número de mensajes escuchados es menos significativo. Esto es debido a que la inestabilidad de la red puede hacer que haya nodos que no hayan escuchado el mensaje aunque se encuentren dentro del área ya cubierta por otras retransmisiones. Una posible solución sería utilizar un mayor umbral *C* pero disminuiría el ahorro de mensajes con el consecuente incremento de la sobrecarga.

Debido a estos problemas, consideramos que el algoritmo tradicional basado en contador puede mejorarse haciendo que cada nodo de la malla (*forwarder*) pueda variar su umbral *C* dependiendo de las condiciones locales de la red. Para que sea posible identificar el estado de la red, necesitamos una métrica adaptable y distribuida que informe al nodo sobre la estabilidad de los enlaces. Si los enlaces son estables, los mensajes que el nodo está escuchando probablemente también estén siendo escuchados por otros que se encuentren en la misma área. A continuación, definiremos la métrica que proponemos y después describiremos nuestra adaptación del algoritmo basado en contador.

3.1 Métrica de Movilidad: Intervalo Modal de la Duración del Enlace

Boleng, Navidi, y Camp [14] usan la métrica de duración del enlace calculada de la siguiente manera: se obtiene la vida media de los enlaces con cada vecino y se establece la duración del enlace como la media de dichos valores. Ellos mostraron que la duración del enlace es un buen indicador del rendimiento del protocolo, ya que es computable en un entorno real y distribuido, a la vez que independiente de cualquier protocolo específico.

Esta métrica cumple muchos de nuestros requisitos: es distribuida y refleja el estado de la red. Sin embargo, si un pequeño conjunto de nodos tiene un comportamiento radicalmente diferente del resto, la media puede no reflejar el comportamiento de la mayoría de los nodos. Nosotros deseamos una métrica que refleje el comportamiento de la mayoría de los nodos *forwarder*, escondiendo la distorsión producida por aquellos nodos con un comportamiento muy diferente del resto. Por tanto, en vez de la media, nosotros hemos adoptado un intervalo modal de la vida media de los enlaces con los vecinos.

La forma en que calculamos esta métrica es la siguiente: primero dividimos la recta real en intervalos disjuntos (por ejemplo en Fig. 3, el eje x ha sido dividido en 3 intervalos). Después, durante un periodo de tiempo *T*, cada nodo calcula la duración media del enlace DE_f para cada uno de sus vecinos *forwarder*. Con el fin de establecer cuándo los enlaces están levantados, el periodo *T* se divide en *k* ranuras cuya duración *t* ha de ser lo suficiente larga para permitir que el nodo reciba al menos un mensaje de cada vecino (algo que ocurre cada *flooding timeout* de los mensajes de control). Entonces la duración media del enlace para cada nodo *forwarder* DE_f se define como:

- La función $h_f(i)$ determina si el enlace con el nodo *f* está levantado en la ranura *i*.

$$h_f(i) = \begin{cases} 1 \; si \;\; escucho \;\; un \;\; mensaje \, desde \;\; f \\ 0 \;\; en \;\; otro \;\; caso \end{cases}$$

- La función Ch_f calcula las veces que el enlace se levanta durante el periodo T.

$$Ch_f = \begin{cases} \sum_{i=1}^{i=k-1} \overline{h_f(i)} \bullet h_f(i+1) \; si \; h_f(1) = 0 \\ 1 + \sum_{i=1}^{i=k-1} \overline{h_f(i)} \bullet h_f(i+1) \; si \; h_f(1) \neq 0 \end{cases}$$

- Entonces la función DE_f (duración media del enlace con f) se calcula dividiendo el tiempo en el que el enlace con el nodo f está levantado entre el número de veces que el enlace se levanta, durante el periodo T.

$$DE_f = \begin{cases} \frac{\sum_k h_f(i)}{Ch_f} \; si \; Ch_f \neq 0 \\ \sum_k h_f(i) \; si \; Ch_f = 0 \end{cases}$$

En tercer lugar, cada DE_f pertenece a un intervalo del conjunto previamente definido en el paso 1, por tanto el intervalo modal de la duración del enlace *MDE* es aquel intervalo al que pertenece el mayor número de DE_f.

Con el fin de que esta métrica se adapte de forma asintótica a las condiciones cambiantes de la red, el periodo de cálculo T presenta una estructura de ventana. La métrica se computa cada t unidades de tiempo, siendo t la duración de cada ranura ($t=T/k$). Es decir, si calculamos la métrica durante un periodo T que abarca desde el instante m hasta el instante $m+k*t$, en el siguiente cálculo de la métrica, el periodo T abarcará desde el instante $m+t$ hasta el instante $m+(k+1)*t$.

Esta métrica permite al nodo hacerse una idea de qué está sucediendo con la mayoría de sus nodos vecinos, es decir, qué ocurre en su área. Por ejemplo, si el valor de *MDE* es alto, la red de su alrededor es básicamente estable aunque haya un número pequeño de nodos con enlaces inestables.

3.2 Método Basado en Contador Adaptable a la Movilidad

El algoritmo basado en contador adaptable a la movilidad es una variación del algoritmo de contador básico. En él hay dos procesos trabajando concurrentemente: el primero calcula el valor del umbral C como función de *MDE*, *C=CC(MDE)*, mientras que el segundo aplica el mecanismo basado en contador. Fig. 2 muestra cómo funciona al algoritmo basado en contador adaptable a la movilidad.

Nosotros hemos elegido una función *C=CC(MDE)* con la forma de Fig. 3 basándonos en las siguientes heurísticas:

- Si el valor de *MDE* es bajo, el nodo sólo establece enlaces de corta duración. En este caso el nodo sólo retransmitirá el mensaje si lo ha escuchado muy pocas veces. Esto es porque consideramos que en una red densa (donde escuchará el mensaje muchas veces), la retransmisión probablemente resultaría innecesaria, pero si la red es dispersa puede ser que no haya otros nodos capaces de retransmitir el mensaje aparte del actual.
- Si el valor de *MDE* es medio, el nodo está en una red donde puede establecer enlaces de una duración moderada. En este caso un valor medio de C ayuda a cubrir el área.
- Si el valor de *MDE* es alto, la red es básicamente estable: hay pocas roturas de enlaces. Por tanto un umbral bajo será suficiente.

```
buclePrincipal()
{
  msj=recibirMensaje();
  si (msj.esDuplicado==falso)
  {
       msj.estableceContador(1);
       msj.estableceTemporizador(aleatorio(0,..tmax));
       msj.iniciaTemporizador();
  }
  si no
  {
       si (msj.temporizadorFinalizado()==falso)
            msj.incrementaContador();
  }
...
}

/* Manejador de eventos llamado cuando
el temporizador expira para el mensaje 'msj' */

temporizadorExpirado(msj)
{
  C=CC(MDE);
  si (msj.valorContador==C)
       msj.descartar();
  si no
       msj.retransmitir();
}
```

Figura 2: Pseudocódigo que describe el método basado en contador adaptable a la movilidad

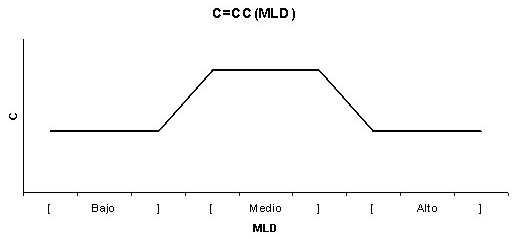


Figura 3: Forma de la función *C=CC(MDE)*

4 Resultados de las Simulaciones

Los mecanismos de reducción de la sobrecarga de datos han sido aplicados a un protocolo de enrutamiento multicast para determinar su rendimiento. El protocolo que hemos escogido ha sido ODMRP. ODMRP no sólo es un protocolo bien conocido, sino que además presenta, como puede observarse en [4], un buen rendimiento comparado con otros protocolos de enrutamiento multicast.

ODMRP ha sido modificado para dotarlo con los mecanismos de reducción de la sobrecarga: todos los nodos *forwarder* ejecutan alguno de los algoritmos basados en contador antes de reenviar un mensaje de datos. Además, cuando se utilizan procedimientos adaptables a la movilidad, los nodos *forwarder* ejecutan un proceso para establecer el umbral *C*. Con este fin, los intervalos que hemos considerado para categorizar *MDE* son los siguientes: si *MDE<=21* segundos lo consideraremos un valor bajo y estableceremos *C=2*, si *21< MDE<=75* segundos consideraremos *MDE* medio y estableceremos *C=3*, para *MDE>75* consideraremos *MDE* alto y estableceremos *C=2*.

El rango que delimita los enlaces menos estables se define tomando como base [14], donde se muestra que el hecho de que los enlaces posean una vida media menor de 15-20 segundos lleva a una red inestable donde el rendimiento de los protocolos decae bruscamente. Respecto al rango superior, estimamos que un enlace cuya vida sea superior a 75 segundos es realmente estable en este entorno. El periodo de tiempo *T* durante el que se calcula *MDE* es de 90 segundos, ya que podemos capturar en la ventana tanto los enlaces con una vida corta (menor de 20 segundos) como aquellos con una vida larga (mayor de 75 segundos). *MDE* se recalcula cada 3 segundos ya que este es el periodo de inundación de la red. Hemos elegido *C=3* como el valor medio, ya que según [7] la "cobertura adicional prevista" en ese caso está sobre el 10%. Las simulaciones se han llevado a cabo utilizando NS-2 [15] en su versión 2.1b8, con las extensiones multicast desarrolladas por el proyecto *Rice University Monarch* [16] que incluyen la implementación de ODMRP. El escenario simulado está formado por 100 nodos móviles distribuidos aleatoriamente en un área de 1600x1200m. La capacidad del canal de radio es de 2Mb/s para cada nodo, los cuales tienen un radio de cobertura de 250m. A nivel de enlace se ha utilizado IEEE 802.11b.

Cada una de las variantes del algoritmo basado en contador simuladas se ha evaluado sobre el mismo conjunto generado previamente. Éste consiste en 400 diferentes escenarios, donde se ha variado la velocidad de los nodos y el volumen de tráfico transmitido. El modelo de movilidad ha sido Gauss-Markov [17]. La velocidad se actualiza cada 10 segundos, las desviaciones típicas de velocidad y ángulo son 0.1 y $\pi/8$ respectivamente, y la velocidad máxima es 0, 5, 10, 15 y 20m/s según el escenario. Para obtener diferentes cargas de tráfico se ha utilizado un mismo grupo multicast variando el número de fuentes CBR (1, 2 y 4) así como el de receptores (5, 15 y 30).

4.1 Métricas de Evaluación Utilizadas

Para comprobar la eficacia de los métodos propuestos hemos utilizado las siguientes métricas:

- Proporción de Paquetes Entregados (*Packet Delivery Ratio* o PDR): se define como el número de paquetes de datos correctamente entregados dividido entre el número de paquetes de datos generados por las fuentes.

- Sobrecarga Normalizada (*Normalized Overhead)*: se define como la suma del total de los paquetes (control más datos) enviados y reenviados dividido entre el número total de paquetes de datos entregados con éxito.

- Eficiencia del Reenvío (*Forwarding Efficiency* o FEF): se define como el número medio de veces que un paquete de datos multicast es reenviado antes de llegar a su destino. Esta métrica representa la eficiencia de la estructura de reenvío subyacente.

- Retardo Medio (*Average Delivery Delay*). Para cada receptor, se computa la media del retardo de todos los paquetes que recibe. Luego el Retardo Medio (global) se calcula promediando todas estas medias.

4.2 Análisis de los Resultados

Hemos simulado tres variantes del método de reducción de la sobrecarga basado en contador. Dos con un umbral fijo de *C=2* y *C=3* y una tercera con nuestro contador adaptable a la movilidad. Estas tres variantes han sido comparadas con el protocolo ODMRP sin ninguna medida de reducción de la sobrecarga. Fig. 4, Fig. 5 y Fig. 6 muestran los resultados obtenidos en función de la velocidad máxima de los nodos. Por la forma de construirse la malla, al incrementar el número de fuentes aumenta tanto en el tráfico emitido como la densidad de nodos *forwarder*.

Fig. 4a, 5a y 6a muestran el PDR en función de la velocidad máxima de los nodos. En todas las simulaciones el PDR ofrecido por *C=2* es insuficiente y se encuentra casi siempre por debajo del 95%. Esto es debido a que el número de retransmisiones es insuficiente. En general los métodos basados en contador ofrecen peor PDR que ODMRP en redes dispersas. Esto es debido a que la malla no tiene suficiente redundancia; la situación empeora al aumentar la velocidad, ya que la rotura de enlaces es mayor y no hay caminos alternativos. Conforme la red se vuelve más densa, los métodos basados en contador ofrecen un mejor PDR obteniendo resultados similares (a veces superiores) a ODMRP. En el caso de *C=3* la velocidad sólo causa una ligera caída. Sin embargo, nuestro método adaptable se acomoda a mayores velocidades y obtiene mejor PDR cuando la velocidad es mayor de 10m/s.

Fig. 4b, 5b y 6b muestran la sobrecarga normalizada. En todos los casos, el uso de métodos basados en contador reduce considerablemente la

sobrecarga. La variante que funciona mejor es *C=2*, pero como ya se ha dicho, su rendimiento es demasiado bajo. *C=3* es el que proporciona un ahorro menor (entre un 17% y un 42% comparado con ODMRP). Este ahorro es directamente proporcional a la redundancia de la red, pero casi no se ve afectado por el aumento de la velocidad. Nuestro método adaptable proporciona un ahorro mayor que *C=3* pero menor que *C=2* (entre un 27% y un 56% comparado con ODMRP). También proporciona una mayor reducción de la sobrecarga cuando la malla es densa pero además, el ahorro también varía con la movilidad: el ahorro es mayor en entornos de baja movilidad donde los enlaces son estables. Esto es porque nuestro método detecta la estabilización de la red, pasando entonces a reducir el número de mensajes transmitidos.

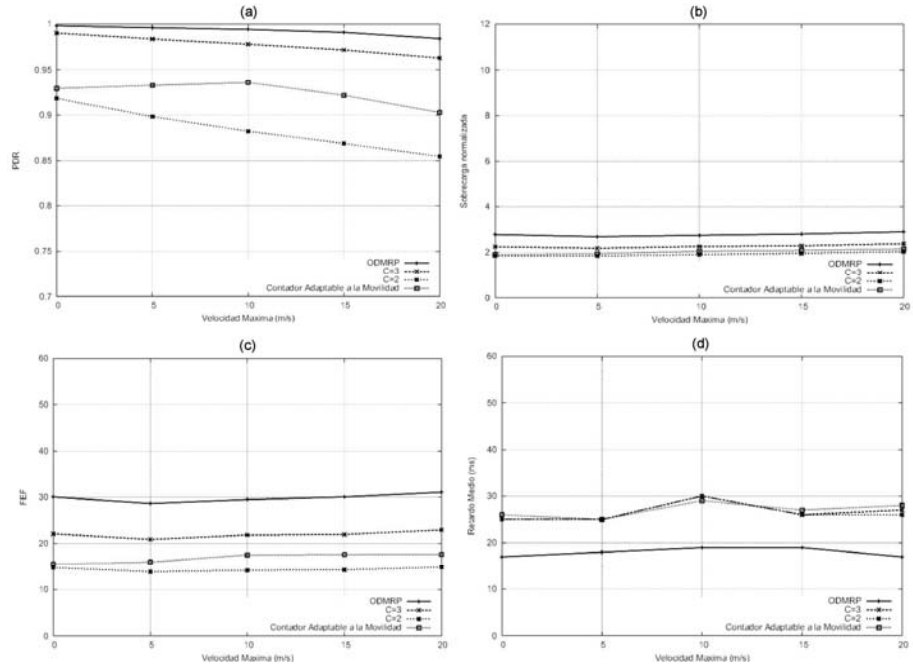


Figura 4: Resultados con 1 fuente y 15 receptores

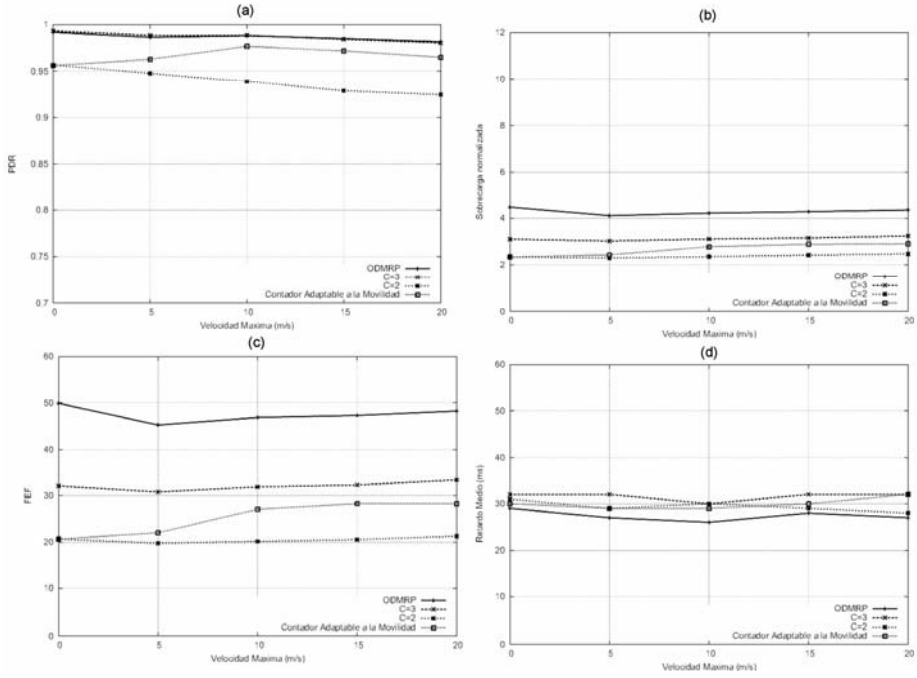


Figura 5: Resultados con 2 fuentes y 15 receptores

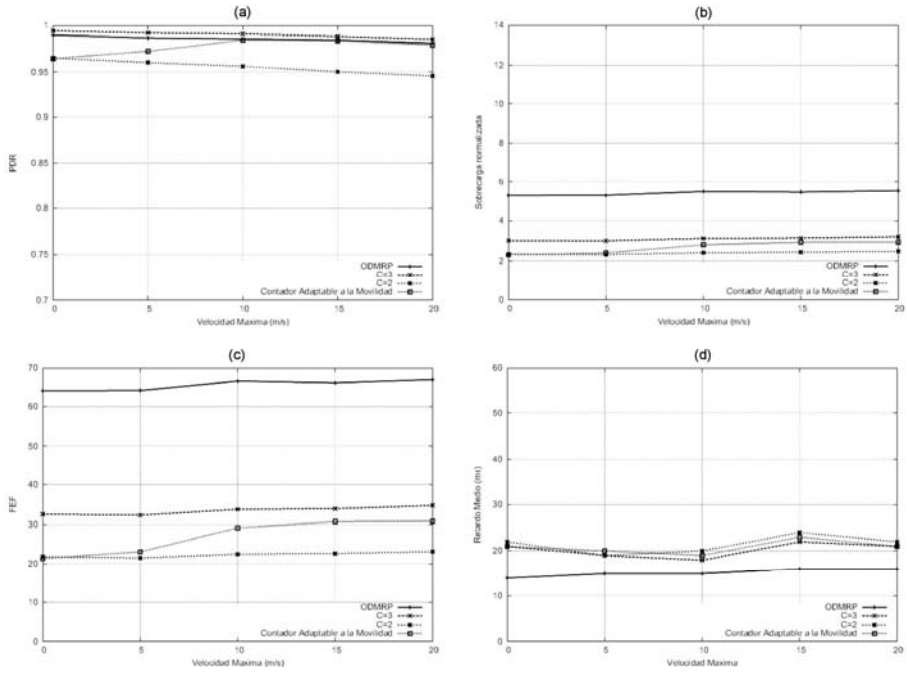


Figura 6: Resultados con 4 fuentes y 15 receptores

Las Fig. 4c, 5c y 6c muestran el FEF. En ellas se observa que, cuando se utilizan métodos basados en contador, el número de paquetes que han de ser reenviados experimenta un importante descenso comparado con ODMRP. Ello es debido a que estos métodos evitan el reenvío innecesario de paquetes. Los resultados son similares a los de la sobrecarga, por tanto nuestro método proporciona menor FEF que *C=3* pero mayor que *C=2*.

Las Fig. 4d, 5d y 6d muestran el retardo medio en milisegundos. Todos los métodos basados en contador presentan un retardo medio similar y mayor que ODMRP. Esto es debido a que ODMRP sigue un enfoque *shortest path tree* pero, al usar un sistema de reducción de sobrecarga basado en contador, el camino que sigue el mensaje no tiene por qué ser el más corto ya que hay nodos *forwarder* que se inhiben de retransmitir.

En resumen, estas gráficas muestran que los métodos de inhibición de reenvío basados en contador pueden ser aplicados con éxito en protocolos de encaminamiento multicast basados en malla. En general, todas las versiones simuladas disminuyen la sobrecarga y el reenvío se hace con una mayor eficiencia. También se deduce de los resultados de las simulaciones que este método funciona mejor en entornos donde la densidad de nodos *forwarder* es grande. Con respecto a las variaciones que usan un contador fijo, con *C=3* se obtiene una importante reducción de la sobrecarga a la vez que se consigue un PDR similar a ODMRP. Sin embargo el umbral *C=2* se ha revelado insuficiente a pesar de ser el que más reducción de la sobrecarga produce (llega a alcanzar una reducción del 60%).

Con respecto a nuestro método adaptable a la movilidad, su rendimiento puede situarse entre ambos umbrales: normalmente ofrece un PDR aceptable, cercano al ofrecido por *C=3* pero con mejor FEF y mayor reducción de la sobrecarga. La curva de la sobrecarga está cercana a *C=2* cuando la movilidad es baja (hay pocas roturas de enlaces) y se aproxima a *C=3* al crecer la movilidad (aumenta la rotura de enlaces). El mejor rendimiento se ofrece cuando la movilidad es baja pero la red no es completamente estática, ya que ofreciendo un PDR mucho mejor que *C=2*, la disminución de la sobrecarga es prácticamente la misma.

5 Conclusiones

Los protocolos de encaminamiento basados en malla disponen de caminos alternativos que, si bien garantizan su robustez, introducen una considerable redundancia en la transmisión de los mensajes de datos. Esta redundancia ha de ser reducida sin que por ello la robustez del protocolo se vea aceptada. Con este fin, hemos hecho que los nodos de la malla decidan si retransmitir o no utilizando el algoritmo tradicional basado en contador. Este método se ha revelado como una buena solución para reducir la sobrecarga de datos manteniendo la robustez. Sin embargo, este algoritmo se basa en un valor umbral fijo: si el umbral es bajo se consigue un gran ahorro de mensajes pero el rendimiento del protocolo baja, y viceversa. Por esta razón proponemos una variación de este algoritmo. El mecanismo basado en contador adaptable a la

movilidad obtiene un buen rendimiento como sucede con un valor umbral alto pero ofrece una mejor eficiencia.

Utilizando el mecanismo basado en contador adaptable a la movilidad cada nodo *forwarder* cambia su valor umbral de acuerdo con la estabilidad que presenta la red de su entorno. Para que el nodo pueda conocer el estado de red, nuestro método utiliza una nueva métrica, el intervalo modal de la duración del enlace, que proporciona al nodo una idea del comportamiento de la mayoría de los nodos que se encuentran a su alrededor.

Utilizando nuestro el mecanismo basado en contador adaptable a la movilidad no se poda la malla y en consecuencia se conserva una estructura robusta: mantenemos el rendimiento del protocolo y la sobrecarga puede llegar a reducirse en un 56%.

Agradecimientos

Este documento ha sido financiado parcialmente por el MCYT español a través del proyecto SAM (TIC2002-04531-C04) y el programa "Ramón y Cajal".

Referencias

[1] C. de Morais Cordeiro, H. Gossain, D. P. Agrawal. "Multicast over Wireless Mobile Ad Hoc Networks: Present and Future Directions". IEEE Network, pp. 52–59, vol. 17, no. 1. (2003).

[2] L. Ji, M. S. Corson. "Differential Destination Multicast — A MANET Multicast Routing Protocol for Small Groups". Proceedings of the INFOCOM, 1192–2002, 2001.

[3] E. M. Royer, C. E. Perkins. "Multicast Operation of the Ad Hoc On-Demand Distance Vector Routing Protocol". Proceedings of the Int'l Conf. Mobile Computing and Networking (MOBICOM). 207–218, Agosto 1999.

[4] S. J. Lee, M. Gerla, C.C. Chiang. "On Demand Multicast Routing Protocol". Proceedings of the IEEE WCNC'99. 1298-1302, Septiembre 1999.

[5] A. McAuley, M.-Kang Liu, R. Talpade E. Bommaiah. "AMRoute: Adhoc Multicast Routing Protocol," Internet draft, 1998.

[6] P. M. Ruiz, A. F. Gómez-Skarmeta. "Mobility-Aware Mesh Construction Algorithm for Low Data Overhead in Multicast Ad Hoc Routing". Journal of Communications and Networks (JNC), pp. 331-342, vol. 6, no. 4. (Diciembre 2004).

[7] S.-Y. Ni, Y.-C. Tseng, Y.-S. Chen, J.-P. Sheu, "The broadcast storm problem in a mobile ad hoc network". Proceeding of the Int'l Conf. Mobile Computing and Networking (MOBICOM), 151-162, Agosto 1999.

[8] Y. Yi, M. Gerla, T. J. Kwon. "Efficient flooding in ad hoc networks: a comparative performance study". IEEE International Conference of Communications, pp. 1059-1063, vol. 26, no. 1. (Mayo 2003)

[9] H. Lim, C. Kim. "Multicast tree construction and flooding in wireless ad hoc networks". Proceedings of the 3rd ACM international workshop on Modelling, analysis and simulation of wireless and mobile systems, 61-68, 2000.

[10] A. Qayyum, L. Viennont, A. Laouiti. "Multipoint realying: An efficient technique for flooding in wireless ad hoc networks". INRIA report, Marzo 2000.

[11] Karthikeyan Chandrashekar, John S. Baras. "Multicast Routing in Mobile Ad Hoc Networks Using Source Grouped Flooding". ISR Technical Report CSHCN 2003-12. 2003.

[12] J. Cartigny, D. Simplot. "Border Node Retransmission Based Probabilistic Broadcast Protocols in Ad Hoc Networks". Proceedings of the 36th Hawaii International Conference on System Sciences 2003.

[13] Y. Tesng, S. Ni, E. Shih. "Adaptive Approaches to Relieving Broadcast Storms in a Wireless Multihop Mobile Ad Hoc Network". IEEE Transactions On Computers, pp 545-557, vol. 52, no. 5, May 2003.

[14] J. Boleng, W. Navidi, T. Camp. "Metrics to Enable Protocols for Mobile Ad Hoc Networks". Proceedings of the International Conference on Wireless Networks (ICWN '02), pp.293-298, 2002.

[15] The Network Simulator Ns-2: http://www.isi.edu/nsnam/ns/.

[16] The Rice University Monarch Project: http://www.monarch.cs.rice.edu/.

[17] T. Camp, J. Boleng, V. Davies. "A Survey of Mobility Models for Ad Hoc Network Research". Wireless Communications & Mobile Computing (WCMC): Special issue on Mobile Ad Hoc Networking: Research, Trends and Applications, pp 483-502,vol. 2, no. 5 (2002).

Caracterización estadística de un servicio real de video bajo demanda

D. Melendi, V.G. García, M. Vilas, R. García, X.G. Pañeda, I. Rodríguez
Departamento de Informática. Universidad de Oviedo
Edificio Departamental 1. Campus Universitario de Viesques SN
33204 – Gijón (Asturias)
Teléfono: 985 18 25 96 Fax: 985 18 19 86
E-mail: {melendidavid, victor, vilasmanuel, garciaroberto, xabiel, rodriguezisabel}@uniovi.es

***Abstract**. The aim of this paper is to perform the analysis, characterization and modelling of the www.lne.es video-on-demand service (**LNE TV**). Its principal special characteristic is the wide range of subjects and lengths of the offered contents. We have designed a method to develop lab experiments for audio/video services. Elements about user behaviour have been analyzed such as session analysis, delivered time, pause distribution, jumps length, etc. Finally, an accurate workload characterization has been done, having all the information necessary to construct a model of the audio/video on demand service. This study has been performed thanks to an access log database with more than 150,000 reproductions of almost 900 videos stored over a period of 4 years. The conclusions of the study are essential to improve the service configuration and content selection, which can help administrators to predict future situations and avoid performance problems.*

1 Introducción

Actualmente, el número de accesos web provoca grandes transferencias de información a través de las redes que, unido al aumento del ancho de banda en los accesos de los usuarios, ha provocado la aparición de nuevos servicios, como el audio/video en Internet. En el audio/video bajo demanda (VoD), el usuario solicita la información en un instante de tiempo y el servidor la entrega al usuario solicitante de forma exclusiva. Este sistema, basado en la tecnología de *streaming*, permite que el usuario interactúe con el servidor, siendo el comportamiento similar al de un reproductor de video.

En este artículo se abordará el desarrollo de un modelo de un servicio real de VoD. Los análisis realizados se han basado en el servicio de video bajo demanda del periódico digital La Nueva España (www.lne.es, LNE TV), que, según el índice OJD (Oficina de Justificación de la Difusión), ocupa la octava posición en el ranking de los periódicos digitales en España. Sus características especiales, como son el amplio abanico de temas ofrecidos (noticias, música, cultura, turismo, naturaleza, deportes, etc) y duraciones de los contenidos (desde 2 minutos hasta 2 horas), hacen que éste servicio sea interesante como caso de estudio para la evaluación de los recursos y prestaciones del servicio en diferentes tipos de redes. La realización del modelo ha sido posible gracias a la información extraída de los *logs* almacenados en una base de datos con más de 150.000 reproducciones de más de 900 videos durante la monitorización del servicio desde su aparición en el año 2001 hasta la actualidad.

Para el modelado del servicio de VoD se ha descrito, en primer lugar, el método que especifica los pasos necesarios para la elaboración de experimentos de laboratorio para el servicio considerado. A continuación se ha analizado el comportamiento del usuario, teniendo en cuenta la información disponible de la puesta en funcionamiento del servicio real. Finalmente, el análisis del volumen y tipo de información intercambiada entre los dispositivos que integran la arquitectura del sistema ha posibilitado caracterizar la carga de tráfico originada. Los resultados de esta etapa de análisis y caracterización han permitido definir un modelo de video bajo demanda que refleja la funcionalidad de los dispositivos implicados. Las conclusiones del estudio son esenciales para mejorar la configuración del servicio y la selección de contenidos.

La aportación principal de este trabajo ha sido la elaboración de un modelo basado en el comportamiento real un servicio con información variada, duraciones de los contenidos que van desde los 2 minutos hasta casi 2 horas y que se ofertan a todo tipo de usuarios. A diferencia de otros trabajos en el mismo campo, donde el estudio se centra únicamente en entornos universitarios o centros de investigación, la variedad de los contenidos y destinatarios no está limitada a una temática y destinatarios específicos, lo que condicionaría la actitud de los posibles usuarios del sistema. Además, la gran cantidad de accesos, videos y datos analizados hacen que los resultados obtenidos puedan considerarse estadísticamente válidos.

El resto del artículo está organizado de la siguiente manera. En el apartado 2 se hace una breve revisión de los trabajos previos en este campo. En el apartado 3 se describe el método diseñado para el desarrollo de experimentos de laboratorio con servicios de audio/video. El apartado 4 está dedicado al análisis del comportamiento del usuario, basándose en la monitorización del servicio real que se ha

implementado, mientras que el apartado 5 muestra los resultados del análisis y la caracterización del tráfico generado por los diferentes dispositivos involucrados en el funcionamiento del servicio. El apartado 6 presenta un modelo del cliente y del servidor del servicio de VoD. Las conclusiones serán abordadas en la última sección.

2 Trabajos previos

El análisis de servicios de VoD es un campo reciente en el mundo de la investigación. Los servicios de VoD no han alcanzado todavía un despliegue importante en Internet, de forma que los estudios referentes a su análisis todavía no son abundantes. A pesar de ello, se han publicado algunos trabajos interesantes sobre el análisis de servicios de *streaming* durante los últimos años. Estos trabajos analizan diversos aspectos acerca del comportamiento del usuario, calidad del servicio y popularidad de los contenidos.

En la documentación revisada hay varios artículos que centran su estudio en el comportamiento del usuario, donde se han tenido en cuenta elementos como duración de las sesiones, tiempo de envío de información, interacciones del usuario[1,2]. En [3], el estudio se ha centrado en la conexión del usuario, red de origen del usuario y calidad del video solicitado. La calidad de servicio se ha evaluado en publicaciones como [3,4], en las que el objeto de estudio han sido parámetros como las pérdidas de paquetes, *jitter*, retardos y la calidad percibida por el usuario del servicio. Otro elemento que ha sido ampliamente analizado es la popularidad [1,2,5,6]. La distribución de las solicitudes de los usuarios entre los diferentes videos ofertados se ha comparado con la distribución de Zipf, calculando el parámetro θ que define a esta ley de Zipf generalizada. Basándose en algunos de los resultados de estos trabajos, han aparecido nuevos estudios y herramientas de simulación del comportamiento del servicio real y evaluación de su rendimiento. Así, en [7] los autores evalúan la capacidad del servidor utilizando la carga obtenida por simulación. El trabajo de Shundong Jin [8] sigue la misma línea, presentando una herramienta para el análisis de servicios de VoD. Por otra parte, se han realizado estudios sobre el diseño de métricas para servicios de video [9,10]. Debido a que la información a transmitir en este tipo de servicios es continua, es necesario desarrollar métricas específicas, que permitan realizar análisis más precisos para evaluar la calidad del servicio y el éxito de su implantación.

3 Definición de experimentos

Basándose en la experiencia en el desarrollo de servicios de audio/video en Internet, se ha diseñado un método para la definición de experimentos de laboratorio. El método presentado se integra dentro de la metodología descrita en [11] para la configuración y análisis de servicios de VoD. En esta metodología, la fase de análisis del servicio está dividida en dos partes, una de las cuales trabaja con los datos capturados a partir del funcionamiento real del servicio y otra que trabaja en base a predicciones. Esta última utiliza modelos de simulación y entornos de emulación para evaluar el rendimiento del sistema en diferentes situaciones. El método descrito está orientado a especificar los pasos necesarios para la realización de experimentos de laboratorio en este tipo de servicios.

3.1 Proceso de aplicación

El proceso de aplicación se muestra en la Fig. 1, donde se aprecian las diferentes fases a considerar. Se comienza con la especificación del servicio, donde el tipo de servicio y sus características deben ser determinados. Se diferenciará entre servicios de sólo audio o audio/video y servicios bajo demanda o servicios en directo, dando lugar a los servicios Jukebox, Radio en Internet, Video bajo demanda e Internet TV. Dependiendo del tipo, los parámetros que lo caracterizan serán diferentes. El segundo paso es la definición de objetivos, indicando qué tipo de información debe proporcionar el experimento. Una vez que se han definido los objetivos, se debe decidir qué tipo de experimento es más adecuado, bien un modelo de simulación o un entorno de pruebas de emulación. La siguiente fase a considerar es la definición del experimento, compuesta por las tareas de definición de carga (comportamiento del usuario y contenidos), recursos y arquitectura, parámetros y establecimiento de valores para cada parámetro del experimento considerado. Una vez que el experimento ha quedado completamente especificado, se procede a la fase de ejecución y análisis de los resultados obtenidos. Si todo el proceso se desarrolla de forma correcta, los resultados obtenidos en los experimentos de laboratorio no diferirán en forma a los recogidos del servicio real, analizándose mediante el conjunto de tests definido en [11].

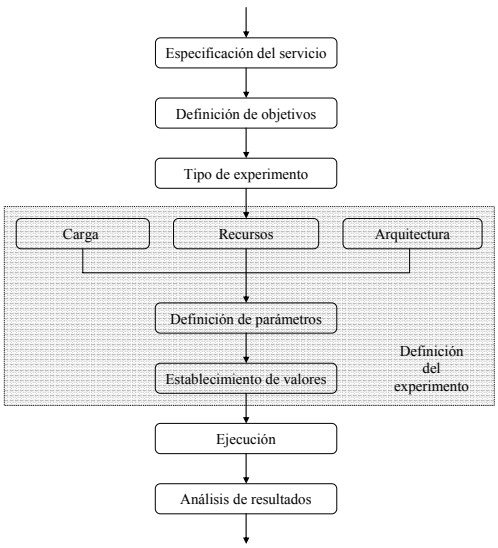


Figura 1: Proceso de aplicación del método

4 Comportamiento del usuario

Se ha realizado el estudio basándose en la monitorización del servicio de video bajo demanda www.lne.es (LNE TV). Se han analizado aspectos del comportamiento del usuario como análisis de la sesión, distribución de las pausas, duración de los saltos en la reproducción, etc. El estudio presenta resultados interesantes sobre las duraciones de los diferentes eventos en la representación, popularidad de los videos, instantes de aparición de las interacciones, etc, que han sido comparados con trabajos anteriores, desarrollados generalmente en entornos educacionales (servicios o usuarios). Se han estimado las distribuciones que mejor se ajustan a los datos del servicio real, permitiendo, a partir de los resultados obtenidos, realizar un modelo que refleja el comportamiento del usuario en este tipo de servicios con información variada.

4.1 Caso de estudio

El estudio se ha desarrollado a partir del servicio de video bajo demanda de La Nueva España Digital (www.lne.es), que cuenta con un número importante de accesos y que ha alcanzado la 8ª posición en el ranking de los periódicos digitales en España. El servicio de video bajo demanda (LNE TV), fue presentado en el año 2001 y ha sido desarrollado por el Departamento de Informática de la Universidad de Oviedo. El número de visitas y el volumen de información han experimentado un importante crecimiento desde su creación hasta hoy. Actualmente, el servicio tiene una gran reputación debido al buen nivel de su producción propia y al amplio rango de temáticas ofertadas.

4.2 Descripción del servicio

La sección multimedia de www.lne.es tiene una arquitectura formada por dos servidores, un servidor de *streaming* y un servidor de análisis (Fig. 2).

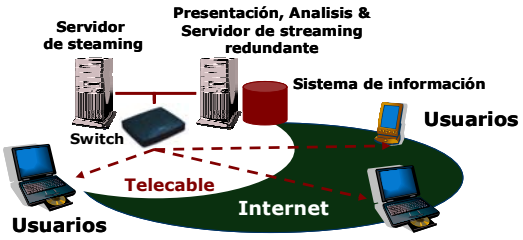


Figura 2: Arquitectura del servicio de VoD de LNE TV

El servidor de *streaming* es el encargado de proporcionar los contenidos de los videos a los usuarios que lo solicitan, mientras que el servidor de análisis soporta la página web de acceso a los videos, el sistema de análisis, actuando, asimismo, como un servidor de *streaming* redundante. La tecnología empleada para transmitir el *stream* de videos es Helix, de RealNetworks [12], de forma que los videos son entregados bajo demanda cuando un abonado realiza una solicitud. El servidor de análisis almacena todos los módulos de la herramienta de análisis [13], lo que incluye la base de datos, el servidor web y los cargadores de datos y analizadores.

4.3. Descripción del contenido

Los contenidos del servicio multimedia se han clasificado en 9 subsecciones de acuerdo a su temática: noticias, música, turismo, ciencia, cine, comedia, ocio, deportes y otros. Las duraciones de los videos abarcan desde 30 segundos hasta casi dos horas. Actualmente, el servicio multimedia de La Nueva España Digital consta de aproximadamente 900 videos disponibles con una calidad de 160kbps. Debido al proceso de análisis y configuración seguido durante el periodo de funcionamiento del servicio, las calidades con las que se ofertan los audio/vídeos han ido variando a lo lago del tiempo. La tecnología utilizada para la codificación de los audio/vídeos ha sido **Surestream** de **RealNetworks,** que permite la variar la calidad en tiempo real con la que se envía el audio/vídeo, si se detectara escasez de recursos en la red de comunicaciones.

4.4 Análisis de la sesión y reproducciones

Con el objetivo de realizar un modelo que represente el comportamiento del usuario, se han evaluado todos los parámetros que lo caracterizan en el servicio de VoD. En la Fig. 3 se muestra de forma esquemática el acceso de los usuarios, donde se hace distinción entre una sesión de usuario y las reproducciones dentro de la sesión. El usuario accede al servicio durante una sesión, que está formada por una o más reproducciones separadas un determinado intervalo de tiempo.

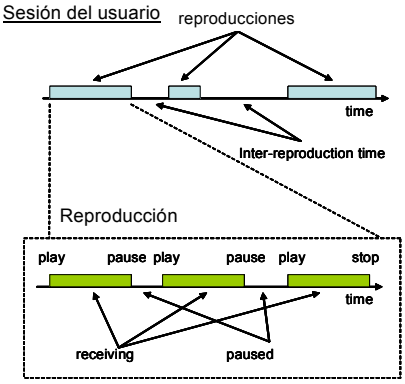


Figura 3: Acceso del usuario en una sesión

Cada reproducción comienza con una interacción *play* y continua hasta el final del audio/video o hasta que se produzca una interacción de *stop*. Durante este tiempo, los usuarios pueden realizar *pause* y *play*, generando periodos de pausa o periodos activos cuando los usuarios están recibiendo la información. Además, se pueden realizar avances y retrocesos rápidos en la reproducción de la información. En el estudio se determinarán las distribuciones teóricas de todos los parámetros que intervienen en el proceso

descrito, de forma que permitan la elaboración de un modelo preciso del comportamiento del usuario en el servicio de video bajo demanda.

4.5 Reproducciones erróneas

La causa fundamental de que se produzcan reproducciones fallidas es la falta del reproductor adecuado en el cliente. A pesar de los mensajes de aviso, por parte del gestor de la Web, de la falta del reproductor y de los detectores *plug-in* instalados, un porcentaje significativo de los accesos falla por esta razón. LNE TV utiliza un script integrado en la página web para detectar si está instalado correctamente RealOnePlayer, necesario para la reproducción, en el PC del cliente. Sin embargo, las reproducciones perdidas mantienen un porcentaje de, aproximadamente, el 15%. Representando el histograma del porcentaje de las reproducciones fallidas, en el 50% de los meses analizados se observa que el 15% de las reproducciones son erróneas, mientras que en el resto de los meses los porcentajes varían entre el 10% y el 20%.

4.6 Información entregada por sesión

Teniendo en cuenta que una reproducción abarca el intervalo de tiempo entre la primera interacción *play* y la interacción *stop* que marca el final, se va a realizar un estudio de la cantidad de información de audio/video entregada (*media delivered*) durante la duración de la reproducción. Ha habido trabajos previos, como [4], en los que se ha detectado la presencia de distribuciones Lognormal y Exponencial para videos cortos (0-5 minutos) y combinaciones de las distribuciones Gamma y Pareto para videos largos (5-50 minutos). En nuestro estudio se han observado también combinaciones de dos distribuciones, como indican los histogramas representados en la Fig. 4.

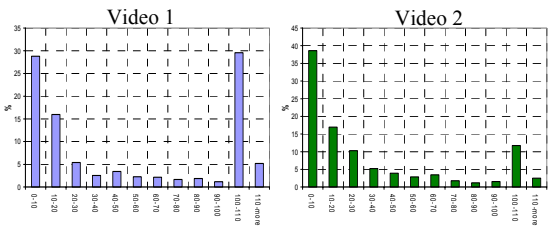


Figura 4: *Media delivered* en una reproducción

En la figura, el video 1 corresponde a un ejemplo de videos cortos, mientras que el video 2 ilustra el comportamiento de los videos largos. Analizando estos histogramas se detecta una primera distribución que abarca desde el intervalo [0-10) hasta el intervalo [90-100). A partir del 100% se aprecia la presencia de una segunda distribución. El peso de cada una de las distribuciones depende del éxito del video analizado, observándose que los videos cortos presentan una mayor presencia de la segunda distribución, lo que indica un mayor interés hacia el video reproducido. En los videos de larga duración el efecto es el contrario, siendo mayor la probabilidad de la presencia de la primera distribución, indicando que, en su mayoría, los usuarios reproducen menos del 100% del video analizado. Utilizando estimadores de máxima similitud (MLE), se ha determinado que la distribución total puede aproximarse mediante la composición de dos distribuciones exponenciales, con parámetros $\mu_1 = 0.16$ y $\mu_2 = 0.06$ para la primera y segunda distribución, respectivamente, en los videos cortos y parámetros $\mu_1 = 0.2$ y $\mu_2 = 0.27$ para los videos largos. La expresión resultante de la combinación de las dos distribuciones viene dada por

$$f(x) = p_1 f_1(x) + (1 - p_1) f_2(x)$$

$$f_1(x) = \begin{cases} \frac{1}{\mu_1} e^{-\frac{x}{\mu_1}} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_2(x) = \begin{cases} \frac{1}{\mu_2} e^{-\frac{x-1}{\mu_2}} & x > 1 \\ 0 & \text{otherwise} \end{cases}$$

siendo μ_1 y μ_2 los parámetros indicados en el párrafo anterior y p_1 la probabilidad de éxito del video considerado.

La presencia de estas dos distribuciones en la caracterización de la duración de las reproducciones es debida a la acción de dos tipos de usuarios que reproducen los vídeos: usuarios con poco interés en el video y usuarios muy interesados en su visualización. La mayoría de las veces los usuarios poco interesados abandonan la reproducción del vídeo durante los primeros segundos e, incluso, su interés decae rápidamente con el tiempo. Sin embargo, los usuarios interesados en el video, no solamente lo visualizan en su totalidad, sino que además realizan saltos hacia atrás para volver a ver determinadas partes del video. Respecto a otros estudios sobre la duración de las reproducciones, en [2] los autores indican que en la mayoría de los casos los usuarios abandonan la reproducción cuando transcurren entre 2.5 y 4.5 minutos, a pesar de la corta duración de la mayoría de las reproducciones. Respecto a [4], la principal diferencia se centra en que LNE TV ofrece programas de noticias y entretenimientos, mientras que [4] ofrece únicamente información educacional. Además, en [4] la visualización de los contenidos era requisito esencial para que los estudiantes superasen la materia, por lo que su comportamiento estaría condicionado por esta circunstancia.

4.7 Pausas en la reproducción

Los análisis realizados sobre las pausas que los usuarios realizan en la reproducción de un vídeo indican que durante una sesión el número de interacciones es bajo. En los vídeos de corta duración, se producen pausas únicamente en el 7% de las reproducciones, resultado lógico debido a la corta duración de los vídeos analizados. Sin embargo, esta tendencia de pocas iteraciones se mantiene también en los videos de larga duración, obteniéndose que el número de reproducciones con pausas no supera el

10% del total. Debido a la naturaleza discreta del número de pausas, se utilizará una distribución discreta para su caracterización.

Además del número de pausas, debe conocerse su duración, ya que este valor determinará el intervalo del periodo *off* en el modelo del usuario. Como se indica en la Fig.5 (vídeos cortos) la caracterización de este aspecto se lleva a cabo mediante una distribución Weibull, tanto para vídeos cortos como para vídeos largos. En los vídeos cortos las pausas tienen una duración media de 55.7 segundos, siendo de 95 segundos para los videos largos.

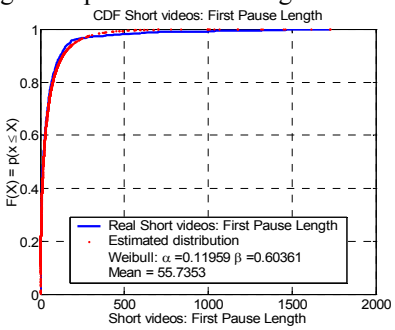


Figura 5: Duración de las pausas en vídeos cortos

4.8 Avances y retrocesos rápidos

A pesar de que los avances y retrocesos rápidos no modifican los periodos *on-off* del modelo de comportamiento del usuario, su caracterización es importante para analizar el rendimiento de cacheado en este tipo de servicios. En el servicio real de LNE TV, el número de avances y retrocesos rápidos en la reproducción se ha modelado siguiendo una distribución Zipf-like (Fig.6), debido a la naturaleza discreta del parámetro a estimar.

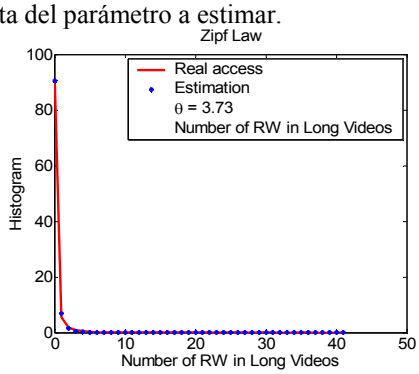


Figura 6: Número de saltos *backward* en videos largos

Todos los posibles saltos y duraciones de los videos quedan caracterizados en la tabla 1.

Tabla 1: Caracterización del número de saltos

Interacción	Videos cortos	Videos Largos
Avance	Zipf (θ = 5.80)	Zipf (θ = 5.22)
Retroceso	Zipf (θ = 4.32)	Zipf (θ = 3.73)

Además del número de saltos en la reproducción debe conocerse su duración. Se ha obtenido la aproximación MLE para la duración de los saltos en el servicio real, mostrando que el mejor ajuste se realiza mediante una distribución Weibull. Los parámetros que caracterizan la distribución Weibull para cada caso analizado se indican en la tabla 2.

Tabla 2: Caracterización de la duración de los saltos

Interacción	Videos cortos	Videos Largos
Avance	α = 0.16827 β = 0.45321	α = 0.10916 β = 0.54129
Retroceso	α = 0.09058 β = 0.47459	α = 0.05959 β = 0.59236

4.9. Características de la sesión

Una sesión está formada por una o varias reproducciones. Al igual que el análisis llevado a cabo en [7], para distinguir una sesión de otra se ha considerado que debe haber un intervalo mínimo de media hora entre dos reproducciones del mismo usuario. La caracterización de una sesión implica la determinación del número de reproducciones que la constituyen y el tiempo entre las reproducciones (*inter-reproduction time* o *reflection time*).

Observando el número de videos en una sesión, se ha estimado que el número de reproducciones en una sesión puede modelarse con una distribución discreta Zipf con parámetro θ = 1.59. El tiempo entre reproducciones ha sido modelado mediante una distribución Weibull con parámetros α = 0.1668 y β=0.5111. El valor medio del tiempo entre reproducciones de una misma sesión es de 78.85 seg.

4.10 Análisis de la popularidad

La popularidad de la información que se ofrece es otro de los factores importantes en el estudio de la carga de un servicio de video bajo demanda. En la documentación revisada sobre el estudio de esta característica, los accesos de los usuarios son estimados mediante una distribución Zipf-like [4,5,6] con diferentes valores del parámetro θ. En [5] el valor estimado para este parámetro es de 0.47, mientras que en [4] los autores empleaban la combinación de dos distribuciones Zipf-like para estimar la popularidad del servicio de video educacional analizado. Si este parámetro es próximo a la unidad, significará que todos los videos ofertados son accedidos menos uniformemente, lo que facilitará el uso de caché con la información. En nuestro estudio se ha detectado que la popularidad de los videos sigue también una distribución Zipf-like, con parámetro θ = 0.667 (Fig.7, izquierda) cuando se tienen en cuenta grandes periodos de tiempo.

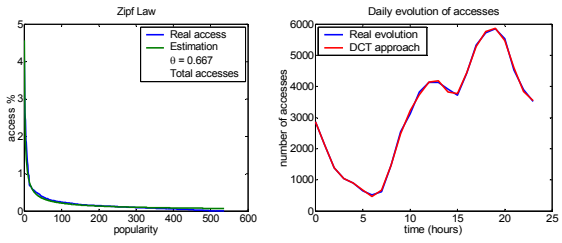


Figura 7: Popularidad del servicio LNE TV desde 2001 hasta 2004 y perfil de acceso diario

4.11 Análisis del acceso diario

El perfil de acceso diario muestra el número de accesos que el servicio presenta a lo largo del día.

Los resultados, mostrados en la Fig.7, con el modelado mediante DCT, indican que durante la noche y primeras horas de la mañana el número de accesos es bajo, incrementándose durante la mañana hasta alcanzar un máximo en la demanda a las 13h. Hay un ligero descenso en los accesos entre las 14h y las 15h, coincidiendo con la hora de la comida, volviendo a aumentar a partir de las 15h hasta las 20h, alcanzando el valor máximo a las 19h. Este perfil se ha mantenido constante durante los cuatro años de funcionamiento del servicio.

Los resultados muestran que, evidentemente, la demanda del servicio se rige por la actividad diaria de los usuarios, marcada claramente por la disponibilidad que fija su horario laboral. Durante el fin de semana, el perfil varía ligeramente respecto a los días laborables. Por un lado, la demanda es ligeramente más baja durante la mañana y el máximo se obtiene a las 23h. Además, el número de accesos al servicio es menor que el resto de los días laborables.

4.12 Modelo de usuario

Se ha diseñado un modelo que refleja el comportamiento del usuario en el servicio de VoD. Para la construcción del modelo se ha considerado una reproducción como elemento básico, de forma que una sesión estará formada por un conjunto de reproducciones. El diagrama que reproduce el comportamiento del usuario se ilustra en la Fig.9. El modelo permite la ejecución de varias reproducciones dentro de una misma sesión. Todos los estados y transiciones en el diagrama han sido caracterizadas con los parámetros indicados en este apartado.

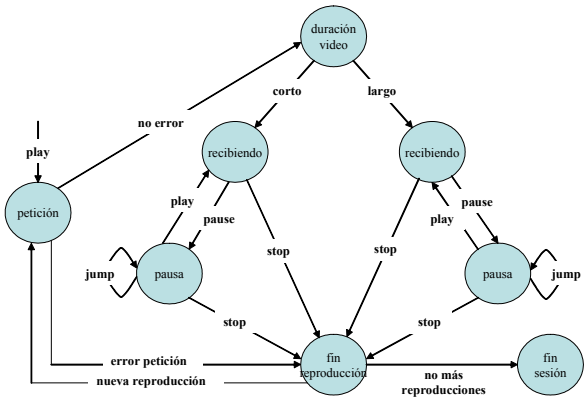


Figura 8: Modelo del comportamiento del usuario

5 Caracterización de la carga

Como última etapa de análisis del servicio se realizará la caracterización de la carga de tráfico generada por los dispositivos que lo integran. Se hará distinción entre el tráfico de diferentes protocolos, dispositivos, flujos de audio y video, así como intervalos de tiempo de carga rápida (*buffering*) y de carga regular. Cuando la información del audio/video se transmite mediante RDT (RealNetworks Data Transport), el cliente RTSP (Real Time Streaming Protocol) establece tres conexiones de red con el servidor RTSP. Una conexión *full-duplex* TCP, que se utiliza para control y negociación de RTSP. Una conexión *simplex* UDP, para RDT, desde el servidor al cliente transporta el audio/video (*media data*), mientras que la segunda conexión *simplex* UDP, en sentido cliente a servidor, se utiliza para sincronización y para solicitudes de reenvío de paquetes perdidos UDP por parte del servidor. Aunque se están haciendo esfuerzos en desarrollar protocolos estándar, muchas aplicaciones comerciales continúan usando protocolos propietarios, lo que dificulta la interpretación de la información capturada, como es el caso del protocolo RDT que se utiliza en el sistema que se va a modelar.

5.1 Videos analizados

Para realizar la caracterización de la carga, se han examinado los videos representados en la tabla 3, donde se indican los parámetros leídos de la cabecera de los ficheros RMFF (Real Media Format File). La nomenclatura utilizada responde al formato

Calidad total(Kbps) – Calidad audio(Kbps) - Frames/seg

Tabla 3: Información resumen de los videos en la cabecera RMFF

Ítem	140-32-20		90-16-15	
	Audio	Video	Audio	Video
Paquetes totales	53191		37649	
Avg packet size (bytes)	465	296	320	288
Max packet size (bytes)	465	666	320	656
Avg bit rate (bps)	32041	107959	16000	74000
Max bit rate (bps)	32041	107959	16000	74000
Preroll (msec)	1857	7483	960	6870
Duration (sec)	998504		998502	
Target Frame Rate	20 frames/sec		15 frames/sec	

Se han elegido para su análisis las calidades 140-32-20 y 90-16-15, debido a que los análisis realizados en [9] indican que estas calidades de audio/video son las que mejor se adaptan a las características de conexión de la mayoría de usuarios que acceden al servicio, bien a través de una red de cable o mediante ADSL. Puede apreciarse cómo el tamaño de los paquetes de audio se mantiene constante, lo cual facilita su identificación dentro de las trazas capturadas para la caracterización del tráfico [14]. Para el tráfico de RealVideo, la caracterización del tamaño de los paquetes es más compleja, al no existir un tamaño definido para este tipo de *stream*.

5.2 Análisis del tráfico a nivel de red

Para llevar a cabo el análisis del tráfico a nivel de red, se han capturado muestras con la herramienta *tcpdump*, de forma que en la monitorización del servicio se han tenido en cuenta los siguientes elementos: dispositivo origen (cliente–servidor), protocolo de transporte (tcp–udp), *timestamp* y tamaño de paquetes. Las medidas realizadas con varias calidades de video, indican que el tráfico UDP de servidor a cliente llega a significar el 95% de los paquetes y el 99% de los bytes intercambiados. Por ello, el consumo de recursos de red de este servicio se

debe, principalmente, a este tráfico UDP, que transporta la información de audio/video, lo que hace necesario un análisis minucioso para caracterizarlo de forma precisa.

Representando el número de paquete UDP originado por el servidor, en función de su estampa de tiempo, y observando su sincronización con el tráfico TCP de control del cliente, se observan 3 situaciones diferentes (Fig.9). Una zona de *preroll* inicial a una mayor velocidad de transferencia, donde los buffers de RealPlayer almacenan la información durante unos segundos antes de comenzar con la reproducción, intervalos de carga regular durante el funcionamiento habitual del sistema y una zona de *buffering*, similar al preroll, donde se necesita una mayor velocidad de transferencia para compensar los efectos de la red. El zoom realizado de la zona de *preroll*, muestra cómo esta zona y la zona de *buffering* se encuentran perfectamente delimitadas por las indicaciones TCP del cliente. La duración de estos intervalos temporales depende de la calidad del video seleccionado y de las condiciones de la red. El tamaño de estos paquetes TCP varía con la calidad seleccionada, entre 211 y 214 bytes.

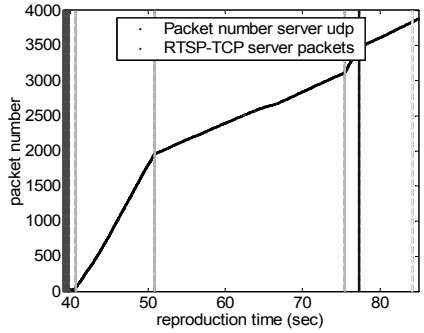


Figura 9: Zonas de preroll, carga regular y buffering

Además de la distinción entre los intervalos de carga regular y carga rápida, un análisis detallado permite diferenciar entre tráfico de audio y tráfico de video. Se ha analizado el comportamiento del tráfico de audio y video en las zonas de *preroll*, carga regular y *buffering*, obteniendo las distribuciones MLE que caracterizan estas situaciones. A modo de ejemplo, la Fig.10 muestra algunos de los resultados obtenidos.

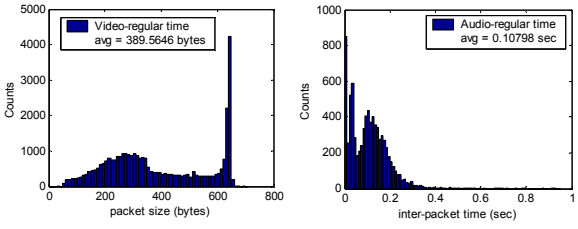


Figura 10: Tamaño de paquetes de video en carga regular y tiempos entre paquetes de audio

6. Modelo del servicio de VoD

Una vez descritos los comportamientos de los diferentes elementos que componen el servicio, se va a desarrollar un modelo que responda al funcionamiento del sistema real.

6.1 Modelo del cliente de VoD

El modelo del cliente del servicio de video bajo demanda debe responder a la funcionalidad descrita en el análisis de comportamiento del usuario, además de ser capaz de establecer las conexiones RTSP y RDT con el servidor para la transferencia de información de control y de *media data*. El diagrama de estados y transiciones diseñado para conseguir la funcionalidad descrita es el mostrado en la Fig.11.

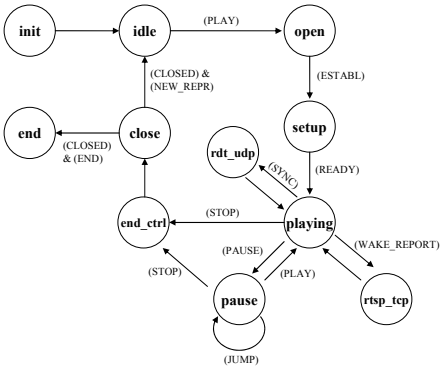


Figura 11: Diagrama de estados del proceso cliente

6.2 Modelo del servidor de VoD

El servidor del servicio de video bajo demanda debe permitir la conexión simultánea con varios clientes que soliciten la visualización de un video (Fig.12).

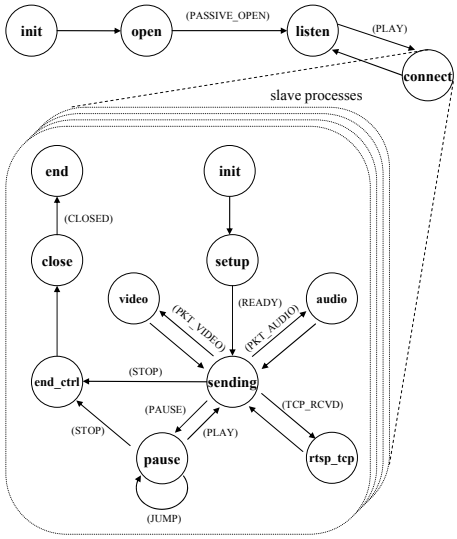


Figura 12: Diagrama de estados del proceso servidor

Además de permitir la conexión RTSP mediante TCP, el servidor enviará los datos del audio/video mediante UDP a los diferentes clientes que lo soliciten y recibirá de estos clientes el tráfico RDT UDP de sincronismo y solicitud de reenvío de paquetes perdidos. Se crean procesos esclavos para la comunicación con cada uno de los clientes que accedan al servicio.

7. Conclusiones

En este trabajo se ha realizado un exhaustivo análisis sobre el comportamiento de los usuarios y la caracterización de la carga de un servicio real de video bajo demanda. El periodo de implantación del servicio analizado, que comprende cuatro años desde su puesta en funcionamiento, la gran cantidad de accesos analizados, más de 150.000 accesos, y los más de 900 vídeos que integran el servicio, son un indicador de la validez de los resultados obtenidos, al contar con un número de muestras significativas para la realización del estudio. Los datos reales de la implantación del servicio han permitido realizar análisis para caracterizar aspectos como la cantidad de información de audio/video transmitida, el número y duración de las pausas, la popularidad de los videos, así como otros aspectos representativos del funcionamiento del servicio. Comparado con otros trabajos previos, este estudio presenta varias ventajas, como son la gran variedad de contenido temático ofrecido a los usuarios, las diferentes duraciones de los videos, la implantación del servicio durante más de cuatro años, además de la variedad de accesos por parte de los usuarios, no estando restringido el acceso a un determinado grupo.

Los resultados han sido esenciales en la definición de un modelo del usuario y de la carga generada que serán integrados en el modelo del servicio completo, lo que permitirá realizar predicciones sobre situaciones diversas, evitando así los problemas de rendimiento que puedan surgir.

Agradecimientos

Esta investigación ha sido financiada por el operador de red Telecable y la Nueva España (proyecto Media XXI) y por el proyecto del Ministerio de Ciencia y Tecnología Integra-Media (Ref: TSI2004-00979).

Referencias

[1] J. M. Almeida, J. Krueger, D. L. Eager, M. K. Vernon: Analysis of Educational Media Server Workloads (NOSSDAV). Port Jefferson, New York, USA. 2001.

[2] M. Chesire, A. Wolman, G. Voelker, H. Lavy: Measurement and Analysis of a Streaming-Media Workload, USENIX Symposium on Internet Technologies and Systems. 2001.

[3] Y. Wang, M. Claypool, Z. Zuo. An Empirical Study of RealVideo Performance Across the Internet. ACM SIGCOMM Internet Measurement Workshop. San Francisco, California, USA. 2001.

[4] D. Loguinov, H. Radha: Measurement Study of Low-bitrate Internet Video Streaming, ACM SIGCOMM Internet Measurement Workshop (IMV). 2001.

[5] L. Chekasova, M. Gupta: Analysis of Enterprise Media Server Workload: Access Patterns, Locality, Content Evolution and Rates of Change, IEEE/ACM Transactions on Networking, 2004.

[6] J. Van Der Merwe, S. Sen, C. Kalmanek: Streaming Video Traffic: Characterization and Network Impact, Seventh International Web Content Caching and Distribution Workshop, Boulder, 2002.

[7] J. R Arias, F. J. Suárez, D. F. García, X. G. Pañeda, V. G. García. Evaluation of Video Server Capacity with Regard to Quality of the Service in Interactive News-On-Demand Systems. Protocols and Systems for Interactive Distributed Multimedia (PROMS-IDMS2002). LNCS 2515. Coimbra, Portugal. 2002.

[8] S. Jin, A. Bestavros: GISMO, A Generator of Internet Streaming Objects and Workloads, ACM SIGMETRICTS. 2001.

[9] X.G Pañeda, D. Melendi, M. García, V. García, R. García, E. Riesgo. Analysis Tool for a Video-on-Demand Service Based on Streaming Technology. 6th IEEE International Conference on High Speed Networks and Multimedia communications. Estoril, Portugal. 2003

[10] J. R. Arias, F. J. Suárez, D. F. García, X. G. Pañeda, V. G. García. A Set of Metrics for Evaluation of Interactive News-on-Demand Systems. ACM International Multimedia Conference. Juan les Pins, France. 2002.

[11] X.G Pañeda, D. Melendi, V. García, R. García, A. Neira. Analysis and Configuration Methodology for Video-on-Demand Services Based on Monitoring Information and Predictions. 6th International Conference on Enterprise Information Systems. Porto, 2004.

[12] RealNetworks. Helix Universal Server. www.realnetworks.com

[13] X. G. Pañeda, D. Melendi, R. Bonis, M. Vilas, I. Rodríguez, R. García. Fesoria, an Integrated Tool for Performance and Content Analysis, SLA Evaluation, Management and Smart Presentation for Video-on-Demand Services. International Conference on E-Business and Telecommunications Networks (ICETE2004). Setubal, 2004.

[14] Mena, A. and Heidemann, J. An Empirical Study of Real Audio Traffic. In *Proceedings of the IEEE Infocom*, p. 101-110. Tel-Aviv, Israel, IEEE. March, 2000.

Servicio CORBA A/V Stream Sin Bloqueo Para Entornos Operativos Heterogéneos

Felipe García Sanchez, Antonio Javier García Sanchez, Pablo Pavón Mariño, Josemaría Malgosa Sanahuja
Departmento de Tecnologías de la Información y de las Comunicaciones, Universidad Politécnica de Cartagena
C/Dr. Fleming, s/n (Campus Muralla de Mar), 30202, Cartagena
Teléfono: 968326537 Fax: 968325973
{felipe.garcia, antoniojavier.garcia, pablo.pavon, josem.malgosa}@upct.es

Abstract. *The CORBA A/V Stream service is a specification of the OMG (Object Management Group) for transmission of audio and video flows employing a middleware communication platform. This service is included in several CORBA implementations (i.e. ORBIX, ACE/TAO). However, the utilisation of middleware for video flows transmission is still infrequent. Some reasons lie on the natural transmission delay introduced by CORBA and the CPU overhead caused by compression techniques and different layer processing. These may be minor issues for some type of data transmissions, but not for video. Therefore, it is interesting to study and improve the features of the service, especially for high rate raw video transmission. Raw video transmission is required for different applications including military, medical and surveillance ones. This paper describes the development of a model for high bit-rate video transmission. It is intensively tested, enhancing their strongpoints. Additionally, some statistical results for the metrics of interest are shown and discussed.*

1 Introducción

La especificación *A/V Stream Service* [1] de CORBA fue presentada por la OMG [2] en Junio de 1998 y revisada en Enero de 2000. La inclusión del modelo de *A/V Streaming* en diferentes implementaciones de CORBA (*Common Object Request Broker Architecture*) sigue esta especificación. Su aplicación ha sido estudiada y analizada ampliamente en diferentes campos [3].

Actualmente, la utilización de *middleware* CORBA en aplicaciones convencionales de intercambio de flujos de vídeo comprimido, es todavía inusual. Las razones son que (1) CORBA introduce una cierta complejidad en los procesos de comunicación, especialmente en el proceso de establecimiento de la conexión, y (2) el procesamiento de señalización añadido por CORBA, asociado a las cabeceras GIOP/IIOP (*General/Internet Inter-ORB Protocol*), de control del flujo de vídeo. En aplicaciones con compresión de vídeo, la inclusión del código de codificación y decodificación a la capa *middleware* obstaculiza el manejo de la información en CORBA [4].

En aplicaciones con vídeo comprimido, la carga de procesamiento añadida y el retardo que ésta produce, se vuelven factores clave, sobre todo, en aplicaciones en tiempo real. Estas desventajas pueden no compensar los beneficios de la capa *middleware* en cuanto a Ingeniería del *Software* (independencia de la plataforma y red de comunicaciones, reutilización de aplicaciones, compatibilidad, etc.).

El vídeo bruto (sin comprimir) es hoy en día estrictamente necesario en determinados campos tecnológicos, como los médicos (tele-operación y tele-diagnosis), los militares (periscopio, visor de tanques) [5] [6] o la tele-vigilancia. En estos y otros entornos, la compresión de vídeo es simplemente no admitida.

El escenario de utilización para las aplicaciones de vídeo bruto difiere de las aplicaciones de vídeo convencionales: (1) el ancho de banda de la red disponible debe ser elevado, (2) la carga de procesamiento en las fuentes y destinos de los flujos de vídeo es menor, al eliminar el proceso de compresión/descompresión. (3) Ello redunda en un retraso de procesamiento más determinista. En este nuevo escenario, los límites difieren de los comentados anteriormente:

- La transmisión de información de control y señalización CORBA, pasa a no ser relevante, en comparación con la alta tasa de bit que requiere el vídeo bruto [7].
- La mayor carga de procesamiento de CPU de la capa *middleware* se ve compensada por el procesamiento más sencillo del flujo de vídeo.

Por estas razones, la aplicación del *middleware* CORBA para el desarrollo de aplicaciones de vídeo bruto encuentra un escenario más ventajoso [5][6]. En este artículo presentamos una aplicación basada en la especificación CORBA *A/V Streaming* destinada específicamente para la transmisión de vídeo de elevado ancho de banda. Además, la aplicación permite la transmisión de información distinta al flujo vídeo, evitando sobrecargar la red. Esta aplicación nos permite evaluar el comportamiento del middleware y, especialmente, las interacciones de los distintos componentes entre sí (aplicación, middleware, sistemas

operativos, protocolos de comunicaciones, red, etc.), ofreciendo resultados selectivos de las plataformas convenientes para este tipo de aplicaciones.

Nuestro modelo CORBA se implementa de acuerdo a dos técnicas diferentes de programación. La primera está basada en la ejecución en paralelo, donde cada aplicación está compuesta por diferentes *threads* o hilos de ejecución. La segunda técnica está basada en programación secuencial. Ambas implementaciones se ejecutan sobre equipos de propósito general y han sido testeadas intensivamente. En particular, hemos estudiado la tasa y el retardo de transmisión. La implementación CORBA seleccionada como soporte, ha sido ACE/TAO. Ésta ofrece su propio *A/V Stream Service*, ampliamente documentado [8], facilitando la adaptación de nuestro modelo.

El estudio se desarrolla en una red de área local, que puede simular diferentes escenarios, como un submarino, una sala de operaciones, etc., [9]. Ambas implementaciones funcionan sobre sistemas operativos comerciales Linux o Windows.

El artículo se divide como sigue: el apartado segundo resume los aspectos más intersantes de la especificación de CORBA A/V Streaming, el apartado tercero presenta el método de invocación empleado, dedicando el cuarto y quinto a los modelos previamente indicados. El apartado sexto muestra los resultados de los estudios y el último ofrece las conclusiones y líneas de trabajo futuras.

2 CORBA *A/V Streaming*

Durante muchos años, los investigadores han trabajado en el desarrollo de *middleware* como plataforma de comunicaciones. Un *middleware* es una capa *software* incluida entre la aplicación (comercial, distribución libre, etc.) y el sistema operativo, para simplificar el desarrollo de aplicaciones distribuidas.

Algunos *middlewares* usan la tecnología ORB (*Object Request Broker*) y dentro de ella, CORBA. CORBA soporta interoperabilidad entre sistemas heterogéneos, proporcionando comunicaciones flexibles y constituyendo la base para desarrollar entornos distribuidos. Sin embargo, los requerimientos necesarios para aplicaciones multimedia distribuidas (control de QoS, tiempo real, manejo de los flujos, etc.) hacen que la política tradicional de *request/reply*, no proporcione los resultados deseados.

El servicio *Audio/Video Streaming* de CORBA soporta la transmisión de flujos de vídeo en tiempo real y proporciona un entorno (*framework*) que facilita la creación, transmisión y recepción de sonido, vídeo y datos entre dos o más fuentes en la red (PC, PDA, equipos móviles, etc.).

El servicio CORBA *A/V Stream* aporta dos ventajas interesantes:

- Proporciona flexibilidad, independizando la aplicación *software* del servidor de vídeo y los clientes de la plataforma de comunicaciones. El programador implementa estas aplicaciones sin atender al lenguaje de cada una de las aplicaciones.
- CORBA no añade sus cabeceras GIOP/IIOP al transmitir los paquetes de vídeo. Este tipo de cabeceras sólo son necesarias para el establecimiento de las conexiones y control, pero no para enviar los datos de vídeo.

La descripción de los principales componentes que forman el modelo del servicio CORBA *A/V Stream* [1] se detalla como sigue (Fig. 1):

- El objeto *Stream Interface Control* proporciona un IDL (*Interface Description Language*) para controlar y manejar flujos. Este tipo de información será transmitido por CORBA usando los protocolos GIOP/IIOP en el nivel de transporte.
- *Flow Data Endpoint*, es un objeto creado desde el *EndPoint* [7], se requiere uno para cada uno de los extremos, servidor y cliente (*Sender* y *Receiver*).
- El *Stream Adaptor* [7] es el objeto que recibe o transmite paquetes sobre la red sin usar las invocaciones del middleware. Los protocolos comúnmente usados son UDP, TCP o RTP.

La Universidad de Washington ofrece una implementación avanzada de CORBA con su entorno ACE/TAO [10] que incluye operación en tiempo real y servicios complementarios al servicio *A/V Stream*. Se selecciona entre otros, MICO, ORBIX u ORBacus, porque está bien documentado y ofrece implementaciones para distintos sistemas operativos (S.O.).

El servicio de TAO *A/V Streaming* proporciona una factoría de fuentes multimedia (*MMDevice*). Un componente *MMDevice* [11] encapsula el comportamiento de una fuente multimedia, que puede ser p. e., una cámara de vídeo. Este componente crea a su vez, *endpoints* para las nuevas conexiones de flujos. Cada *Endpoint* consiste en un par de objetos: (a) Un *Virtual Device* (*VDev*) que encapsula los datos necesarios para la conexión, como puede ser el formato de vídeo, tamaño de paquete, etc., y (b) un *StreamEndPoint*. Este *StreamEndPoint* también encapsula parámetros específicos de la conexión como en nuestro caso el nombre del *host* y el número de puerto para el flujo que utiliza UDP (*User Datagram Protocol*). Además, el servicio de TAO ofrece dos políticas de concurrencia que se encargan de crear ambos objetos *VDev* y *StreamEndPoint*:

- *Estrategia basada en concurrencia de procesos* que crea nuevos *Endpoints* en nuevos procesos. Se utiliza

para aplicaciones que generan procesos separados para el control de las comunicaciones.

- *Estrategia reactiva*, que genera cada nuevo objeto en el mismo proceso. Se aplica a procesos que controlan diversos flujos desde la misma aplicación.

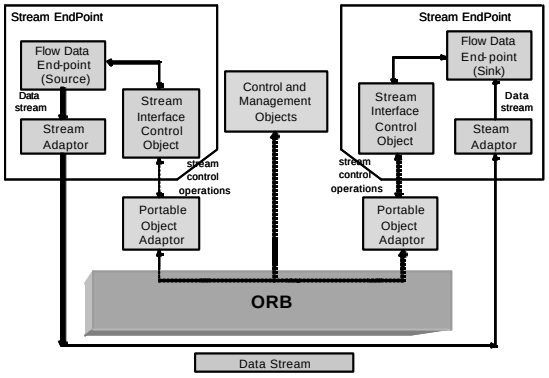


Fig. 1, Servicio *A/V Stream* de CORBA.

Nosotros utilizamos la estrategia reactiva, porque simplifica la implementación del ORB.

3 Modelo de invocación

El modelo de invocación está diseñado bajo dos premisas: la reducción del número de mensajes necesarios entre todos los procesos involucrados, y el desarrollo de un *thread* independiente para la ejecución del ORB. De esta forma, el establecimiento de la conexión se desarrolla en este último *thread*, con las siguientes características:

- Se elimina el bloqueo del ORB (cuando espera la llegada de datos).
- La transmisión de la señal de vídeo no usa el ORB en ningún caso, emplea otro *thread* independiente. Incluso, el vídeo se puede transmitir directamente sobre el protocolo de nivel de transporte (en este caso UDP).
- Se proporciona flexibilidad para añadir nuevas características. Es posible añadir dentro del *thread* del ORB mensajes de control u otro tipo de transmisión de datos.

La reducción del número de mensajes tiene diversos efectos, dependiendo del instante en que actúe. El primero es en el establecimiento de la conexión, compuesto por mensajes GIOP/IIOP que utilizan el protocolo TCP. Para cada nuevo mensaje se genera una nueva conexión TCP entre los extremos. Un mensaje del ORB es suficientemente importante como para garantizar su transmisión y recepción, pero por el contrario se añade retardo. Por esto, la reducción de mensajes GIOP/IIOP es positiva para agilizar la conexión entre procesos.

Otra reducción de mensajes se produce durante la transmisión del vídeo. La transmisión de vídeo bruto requiere información adicional de sincronismo de imágenes y paquetes, etc. En la transmisión de vídeo sin comprimir, se necesita un número importante de cabeceras, separadores de imágenes, etc., que conforman una importante fuente de retardo. El modelo implementa todos ellos en una única cabecera por cada imagen facilitando así la transmisión del flujo sin cabeceras individuales.

3.1 Reducción de mensajes en la conexión

Esta tarea se incluye en el *thread* del ORB y configura la reducción del número de mensajes entre los *StreamEndPoints* y el proceso de control y manejo de servicios. Las principales contribuciones son:

- El proceso de conexión se realiza en el ORB *thread*.
- Los distintos componentes utilizados para la conexión, se reducen a los *StreamEndPoint*, evitando así la sobrecarga de objetos. Se implementa configurando las conexiones extremo a extremo con el protocolo UDP.
- Los datos necesarios tales como la dirección IP y el número de puerto, se transmiten durante la conexión de ambos *StreamEndPoint*.

El procedimiento completo es el siguiente:

1. La aplicación *Sender* diseñada captura la señal de vídeo. Esta señal se envía sin comprimir a nuestra aplicación remota *Receiver*.
2. El proceso *Receiver* (encargado de la captura de vídeo y de su reproducción) se registra en el servicio de TAO de control de objetos (*NameService*) a través del *thread* ORB.
3. La aplicación *Sender* (captura de vídeo y transmisión) encuentra y se conecta al servicio *NameService*.
4. El componente *Receiver* del mismo nombre que la aplicación es reconocida por el *Sender*. Éste recibe la respuesta del *NameService* con la información del *Receiver* lo que arranca, a su vez el *StreamEndPoint* en el *Sender*.
5. El *StreamEndPoint* del *Sender* envía una petición de conexión al *Receiver*.
6. Por último, se conectan los dos *StreamEndPoints* de manera habitual.

El aspecto principal de este establecimiento de conexión es el hecho de que el *Receiver* sea el extremo que se inscriba en el *NameService* en lugar del *Sender*, lo que modifica la estructura clásica en la que el extremo que recibe es el que invoca al transmisor. Además, la información de conexión del *Sender* se transmite en la propia petición de conexión.

Este aspecto particular provoca en el proceso *Sender* mayor complejidad y mayor número de funciones. El *Receiver*, sin embargo, se limita a esperar los datos de vídeo (Fig. 2).

Funciones del proceso *Sender*:

- *Acceso al Sender*. Se debe seleccionar el cliente. Se debe seleccionar el proceso *Receiver* apropiado (de acuerdo, por ejemplo, a una política de prioridades) o rechazar la conexión de algún cliente indeseado.
- *Política de QoS*. El *Sender* toma su propia decisión sobre el tipo de servicio a ofrecer, no existiendo negociación. El *Receiver* puede sugerir los requerimientos, pero es el *Sender* el que dependiendo de la red, del número de procesos, etc., toma las decisiones [12].

Sin embargo, aparecen algunas desventajas:

- La inscripción del *Receiver* fuerza al proceso *Sender* a testear el *NameService* durante la ejecución. El tiempo de establecimiento de la conexión dependerá del tiempo por el que regularmente se realiza el testeo.
- El *StreamEndPoint Sender* no reconoce inmediatamente la finalización del *Receiver*. Ello sólo ocurre cuando el *Sender* testea el *NameService* y comprueba que ya no existe.
- El *Receiver* no puede enviar otros datos (por ejemplo, el control de la cámara) a través del ORB, porque éste no invoca al *Sender* a través de él.

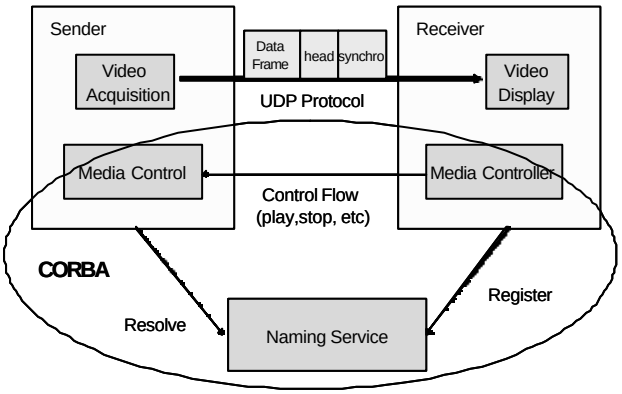


Fig. 2, Esquema de comunicaciones.

Estas limitaciones se producen por la inscripción del *Receiver* en lugar del *Sender*. En la conexión de un único *Sender* y *Receiver*, dichas condiciones no aparecerán porque no es preciso el testeo del *NameService* por el *Sender*. Ello sólo ocurre a partir de un segundo cliente. Para minimizar el efecto de estas desventajas, el *thread* del ORB se encarga de realizar estas funciones, evitando así que afecte al flujo de vídeo.

3.2 Reducción de mensajes externos

La fuente de vídeo bruto genera un flujo, que demanda un elevado ancho de banda, y retarda la transmisión. Los aspectos más importantes son:

- Sobrecarga en las colas de transmisión y recepción.
- Dificultad en el manejo de grandes secciones de memoria.
- Retardo adicional.

De esta forma, un aspecto importante consiste en reducir los datos de transmisión. Algunas técnicas utilizadas son:

- Usar UDP como protocolo de transporte. Las cabeceras son menores y no se incluye reconocimiento.
- Incrementar el tamaño de paquete para el protocolo GIOP/IIOP. Normalmente, en Internet el tamaño del segmento de nivel de transporte es pequeño (512 *bytes*). Sin embargo, nuestras pruebas aconsejan incrementar ese tamaño hasta 16284 *bytes*.
- Limitar la información externa a la cantidad mínima. Se necesita el número de secuencia, sincronismo, etc. Dicha información se envía en el mismo paquete.

Estas reducciones afectan al funcionamiento normal de una aplicación distribuida: control de errores, sincronismo de vídeo, etc.

3.3 El *thread* del ORB

El ORB usa un *thread* independiente para sus funciones. Incluye el establecimiento de la conexión, la fase de la transferencia de datos y el fin de la transmisión.

Todas las funciones se desarrollan adecuadamente para casos de un único *Sender* y un único *Receiver*. Los problemas aparecen cuando el número de *Receivers* aumenta, ya que la necesidad de testear regularmente el *NameService* produce sobrecarga en la CPU y retardo adicional.

Adicionalmente, se debe tener en cuenta otra característica importante. Las aplicaciones suelen requerir otro tipo de información además del flujo de vídeo. Normalmente, se incorporan funciones de control de cámaras, como el zoom) o mensajes de política *request/reply*. Por tanto, se hace necesario una aplicación mixta de flujos y *request/reply*.

De acuerdo con el modelo, no es posible invocar al *Sender* desde el *Receiver*. Por tanto, quedan dos opciones:

1. Usar un nuevo componente para el *Sender*, que pueda ser invocado desde el *Receiver*.

2. Usar el mismo flujo de vídeo para incluir en él la información. Esto supone tener que tratar dichas informaciones de manera separada en ambos extremos, pero elimina sobrecarga.

4. Modelo basado en *multithreads*

Esta implementación se basa en la programación de cada uno de los extremos según diversos *threads*. El proceso *Sender* está compuesto por el *Sender* básico de captura de información, el *StreamEndPoint Sender* y el *thread* dedicado a los mensajes del ORB (conexión, fin de conexión, etc.). El proceso *Receiver* incluye el *thread* original de reproducción de vídeo, el *StreamEndPoint Receiver* y el *thread* para los mensajes del ORB (inscripción en el *NameService*, conexión, etc.). Destacar una vez más que los objetos *VDev* del modelo básico *A/V Stream n*o aparecen, siendo su funcionalidad incluida en los *StreamEndPoint*. La configuración asignada es estática, añadiéndose además una estructura *request/reply* para la transmisión de otro tipo de información.

4.1 La aplicación *Sender*

La aplicación *Sender* se diseña para obtener un buen funcionamiento y racionalizar los recursos del sistema. De esta forma, el proceso original es el encargado de generar el resto de *threads*, (*StreamEndPoint* y ORB). Sus funciones son:

- Conectarse al *NameService*, obteniendo de él la referencia de la aplicación *Receiver*.
- Activar los *threads* auxiliares.
- Invocar los métodos de captura de vídeo, almacenamiento de datos y de lectura y envío del flujo de vídeo.

El *StreamEndPoint Sender* se genera por el método "*Streamctrl*". Ello se realiza cuando se recibe la respuesta del *NameService* y sólo afecta a las conexiones del flujo. La función "*bind_devs*" es la clásica de la propia especificación que a raíz de las dos referencias conecta ambos objetos.

El *thread* del ORB se inicializa en el comienzo de la ejecución, generado por el método "*pthread_create*". Además implementa en él, el testeo del servicio *NameService* en busca de nuevos clientes.

Los métodos implementados en el proceso original son:

1. *Write_data*. Desarrolla la captura de vídeo y almacenado. Cuando está en ejecución impide la actuación de otro método, hasta que se haya producido la captura de una imagen completa.
2. *Stream_Sender*. Lee los datos de vídeo previamente almacenados en memoria, hasta el fin de la imagen. Sólo entonces se permite que el método "*write_data*" pueda volver a actualizar una nueva imagen. Este sistema opera como un mecanismo de exclusión mutua de memoria.

4.2 La aplicación *Receiver*

La aplicación *Receiver* trabaja de acuerdo a la misma metodología que la aplicación *Sender*. Los distintos *threads* están orientados a facilitar la operación de los diferentes métodos y funciones. El proceso original incluye las funciones:

- Crear los *threads* adicionales necesarios.
- Inscribirse en el *NameService*. Esperar hasta que el S*ender* ejecute su petición de conexión.
- Recibir las imágenes (junto con información adicional), y almacenarlas en memoria. Por último, dichas imágenes son leídas y reproducidas.

Los *threads* auxiliares son similares a los *thread* de la aplicación *Sender* y se generan de la misma forma. Sin embargo, el ORB *thread* es más simple, porque su misión consiste en el comienzo del interfaz hacia el *NameService* y la aplicación *Sender*. Una vez realizada la conexión, su tarea se limita a mantenerla. Los métodos incluidos son:

1. *Receive_frame*. Es un método complejo, capaz de separar la información de vídeo y la de control, actuando en consecuencia. Como en el *Sender*, se captura el vídeo y se almacena en memoria.
2. *GetVideoBuffer*. Lee la información de la memoria y les proporciona una función de representación.

5. Modelo secuencial

Este modelo es más sencillo, y consiste en la disposición secuencial de las tareas. La transmisión en CORBA consta de dos partes: captura de vídeo y transmisión de los datos. La adquisición del vídeo usa la herramienta "*Directshow*" (apropiada para Windows) y sus filtros. *DirectShow* es un interfaz aplicación-programa (API) para reproducción, transformación, y captura de una amplia variedad de formatos de datos. En *DirectShow*, se usa el filtro *SampleGrabber*. Este filtro, entre otras funciones, permite acceder al punto de contacto de las imágenes.

Una vez que se obtienen los datos correspondientes a cada imagen, éstos se dividen en paquetes de tamaño fijo, hasta que puedan ser transmitidos por la red de comunicaciones. La máxima ventana de transmisión proporcionada por el modelo es de 16384 *bytes*. Una imagen estándar de 384x288 se divide en 24, 48, 72, 144 ó 288 paquetes, que se corresponde con tamaños de 13824, 6912, 4608, 2304 ó 1152 *bytes*.

Además, para mantener la sincronización entre *Sender* y *Receiver*, se envía un paquete de sincronismo inicial. También, para estudiar los parámetros de interés, es necesario transmitir un número de secuencia y otra

información temporal para el cálculo de los retardos y su variación o *jitter*.

La aplicación *Receiver*, permite la visualización de imágenes enviadas por el *Sender*. El código se divide en dos partes: recepción de datos y presentación.

El método de recepción se invoca por el propio TAO cuando se detectan los datos a la entrada del cliente. El método recibe el nombre de *Receive_frame*. Debe ser definido por el programador, ya que en TAO no realiza ninguna función. Se ha desarrollado de forma que la información quede almacenada hasta que puede ser leída.

Los retardos en los datos a la entrada y salida de las colas de transmisión de los extremos son las causas más probables de bloqueo, pero no las únicas. El propio ORB puede causar bloqueos. El ORB está orientado a eventos, de tal manera que existen funciones y métodos que no se ejecutarán hasta que un evento concreto se produzca. Para evitar esto, el *thread* independiente del ORB se sigue manteniendo para este modelo, de forma que los eventos no puedan bloquear la aplicación completa.

Una vez que se han recibido todos los paquetes de una imagen, se puede proceder a su reproducción. Ello se realiza mediante el método *DrawDibDraw*. Este método utiliza como parámetros el tamaño de la imagen y los datos de vídeo.

6. Estudio de funcionamiento

Para comparar el funcionamiento de los dos modelos (secuencial y *multithread*) a través de una red, se ha realizado un estudio de distintos parámetros como son el *throughput* o utilización, retardo y jitter. Los parámetros [10] se relacionan con sus requisitos de ancho de banda, los retardos y la variación de esos retardos, lo que influye en la adecuada reproducción del vídeo.

El análisis se realiza sobre una red Ethernet de 100Mbps. Se selecciona este tipo de red, porque su protocolo de control de acceso al medio (basado en colisiones), en principio, puede perjudicar a una transmisión en tiempo real. Los equipos utilizados son los mismos (Pentium III a 1 GHz) en ambos casos. Se testean los modelos bajo los S.O. Linux (Red Hat 8.0) y Windows XP.

6.1 Utilización (*Throughput*)

La carga de información útil se puede calcular para cada imagen. Una imagen de 384×288 (tamaño) ×24 (bits por pixel) tiene un tamaño de 2654208 bits. El tráfico generado por el vídeo bruto, es alto y de tasa constante (tráfico CBR) [11].

Sin embargo, una imagen de ese tamaño no se puede transmitir directamente mediante una sola operación, se tiene que dividir en paquetes cuyo tamaño se utiliza como variable de estudio. Este estudio nos proporcionará el diseño más adecuado.

A pesar de elegir el tamaño de paquete como variable independiente, la relación entre imágenes por segundo nos facilita el análisis de la tasa de transmisión. Las distintas tasas reflejan distintos escenarios de funcionamiento, donde las necesidades de vídeo también sufren variaciones. Por ejemplo, no son los mismos requisitos en una tele-operación cuando se está realizando una incisión que cuando se están poniendo unos puntos. Los modelos rechazan las imágenes que lleguen incompletas, que se consideran, a todos los efectos, como información no útil.

Los resultados obtenidos se muestran en la fig. 3. Destacamos los siguientes aspectos:

- Ambos modelos alcanzan un *throughput* similar para la misma tasa de transmisión. Transmitiendo la misma información, sólo se produce un incremento del 3% entre utilizar CORBA o no [12].
- La utilización para cada tasa en imágenes/segundo es proporcional a la información enviada. Así por ejemplo, el valor de la tasa de 25 imágenes por segundo es cinco veces el de 5 imágenes por segundo). De esto se deduce que no hay pérdidas.
- Las variaciones de *throughput* debidas al cambio de tamaño de paquete son razonables pero no despreciables. Especialmente, se deben a desviaciones producidas por cabeceras, paquetes de control, etc.
- El *throughput* es mayor en el modelo secuencial que en el *multithread*. Ello se debe a la mayor tasa de transmisión alcanzada por el modelo secuencial.

La evaluación de ambos modelos demuestra que con el modelo secuencial se puede alcanzar una tasa de transmisión mayor, de hasta 25 imágenes por segundo, mientras que en el modelo *multithread* llega hasta 12 imágenes por segundo sin pérdidas. Esto se debe a que las colas de transmisión para distintos *threads* están limitadas. El paradigma "multitarea" implica que un proceso no pueda tener recursos ilimitados de memoria o CPU. Aunque muchos parámetros son reconfigurables, existen limitaciones de funcionamiento. Por ejemplo,

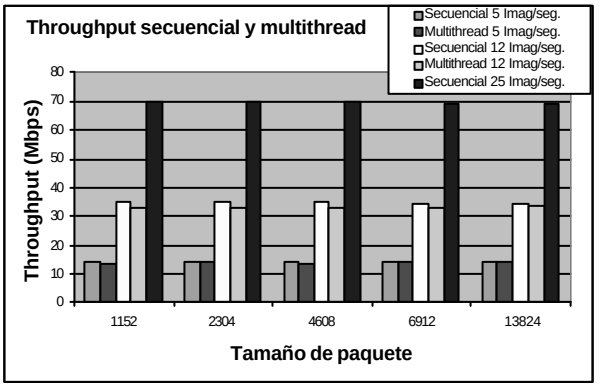


Fig. 3, *Throughput* para los modelos secuencial y *multithread*.

aplicando los tamaños máximos que permite Linux, se comprueba que una aplicación de estas características desborda las colas de transmisión.

En el prototipo secuencial, las fuentes del sistema imponen sus límites, pero no hay más requisitos. Mientras que el sistema disponga de recursos, se pueden utilizar, alcanzándose mayores tasas. Obviamente, en un entorno con varios procesos ejecutándose de manera simultánea, las prestaciones del modelo secuencia se degradan.

6.2 Retardo y *Jitter*

La información necesaria para calcular los retardos se incluye en las cabeceras de sincronismo de cada imagen. El cálculo se realiza en el proceso *Receiver*. Los resultados de retardos y *jitter* se observan en las figuras 4 y 5, extrayéndose las siguientes valoraciones:

- Los retardos son muy similares en ambos modelos para tasas de transmisión bajas. Oscilan alrededor de los 30 mseg., para la tasa de 5 imágenes por segundo.
- Incrementando la tasa de transmisión en imágenes/segundo, los retardos sufren un incremento notable para el modelo basado en *multithreads* (hasta el 100%). Este incremento es menor para el modelo secuencial (entre el 20 y 30%).
- Los valores de *jitter* son mayores para el modelo *multithread*, implementado sobre Linux, especialmente a altas tasas.

Los retardos se producen por los tiempos de espera en la memoria del sistema, perjudicando tanto al *Sender* como al *Receiver*. Estas memorias se diseñan adecuadamente para el *hardware* disponible y en ambos S.O. Esto impide considerar el tamaño de las mismas como razón para este comportamiento.

En el escenario *multithread*, el modelo se optimiza para obtener el mejor rendimiento del funcionamiento *multithread*, implementando sistemas de exclusión mutua, etc. Usando TAO, este proceso resulta más complicado, lo que se traduce en retardo adicional.

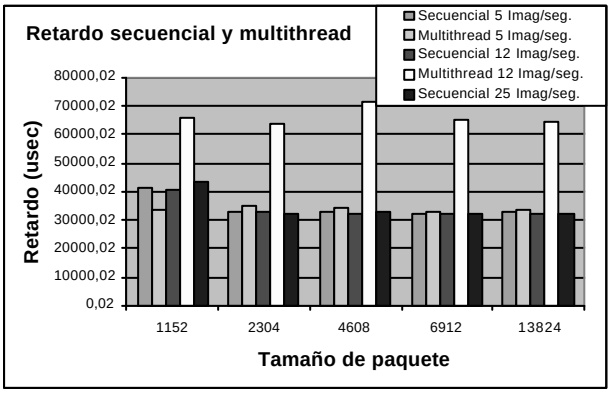


Fig. 4. Retardos en los modelos secuencial y *multithread*.

Además, los retardos para las imágenes son acumulativos, en el sentido de que el retardo en la transmisión de una imagen afecta a la siguiente, probablemente ya capturada, lo cual se traduce en retardo adicional. Este es el factor que hace que el *jitter* también crezca.

Sin embargo, el modelo secuencial no está destinado a sacar provecho de este funcionamiento *multithread*, lo que implica que el proceso de escritura-almacenamiento-lectura sea secuencial. Este proceso es más rápido con un número de procesos bajo.

7. Conclusiones y trabajo futuro

Este artículo presenta un modelo particular del servicio *A/V Stream* de CORBA enfocado a la transmisión de vídeo bruto.

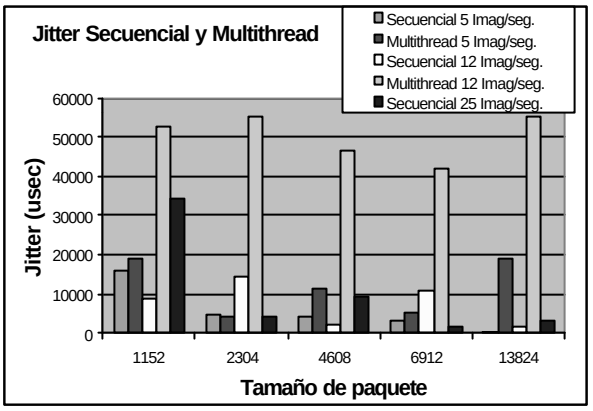


Fig. 5, *Jitter* en los modelos secuencial y *multithread*.

El modelo está basado en minimizar el número de mensajes GIOP y eliminar el bloqueo por el ORB. La principal herramienta es la inversión de la política tradicional donde el cliente es el que invoca al servidor. Aquí, el *Receiver* es el que se inscribe en el *NameService* de TAO y el *Sender* es el que lo invoca.

Se han descrito además detalles de la implementación y se han ofrecido algunas soluciones para los problemas que surgen en el modelo.

El modelo se implementa considerando dos políticas de programación: basada en *multithread* y basada en programación secuencial. Su evaluación comparativa se desarrolla de acuerdo a parámetros como *throughput*, retardo y *jitter*. Dicha evaluación revela lo siguiente:

- El análisis del modelo *multithread* ofrece un comportamiento inesperadamente peor que el ofrecido por el secuencial, especialmente a tasas elevadas. Las características *multithread* unidas a CORBA son la principal causa.
- El modelo secuencial alcanza un comportamiento mejor a tasas elevadas, permitiendo una ejecución

apropiada alcanzando un funcionamiento similar a tiempo real.

- El modelo secuencial ofrece un comportamiento mejor con un menor número de procesos. Conforme éste aumenta, dicho comportamiento se degrada como cabría esperar.

De estas conclusiones, se extraen dos ideas principales:

1. El comportamiento mejor del modelo secuencial en comparación con el de *multithread* cuando el número de procesos es pequeño. Sin embargo, se requieren equipos dedicados.
2. El modelo *multithread* ofrece un servicio estable, independientemente del número de procesos incluidos.

El trabajo futuro incluye:

1. Extender los modelos y estudios a plataformas con sistemas operativos compartidos.
2. Usar diferentes *middlewares* a ACE/TAO, incluyendo otros que no sean CORBA, como JavaRMI, .NET Remoting, etc.
3. Evaluar las mejoras posibles en ACE/TAO para obtener mejores comportamientos con diferentes políticas de programación, políticas *request/reply*, operaciones de *streaming* tradicionales, etc.

Referencias

[1] D. L. Levine, S. Flores-Gaitan & Schmidt, "An Empirical Evaluation of OS EndSystem Support for Real-Time CORBA Object Request Broker". *Multimedia Computing and Networking 2000 (MMCN00)*. San Jose, California, 25-27 January 2000, ISBN: 0-8194-3587-2.

[2] C. D. Gill, F. Kuhns *et al*., "Applying Adaptive Real-time Middleware to Address Grand Challenges of COTS-based Mission-Critical Real-Time Systems", *Proceedings of the 1st International Workshop on Real-Time Mission-Critical Systems: Grand Challenge Problems*, IEEE, Phoenix, Arizona, November 30, 1999. ISBN: 0-7695-1928-8.

[3] M. Atighetchi, P Pal, F. Webber, "Adaptive Use of Network-Centric Mechanisms in Cyber-Defense", *The 6th IEEE International Symposium on Object-oriented Real-time Distributed Computing (ISORC)*, pp. 183-192, Hakodate, Hokkaido, Japan, May 14-16, 2003.

[4] M. Hemy *et al*., "Evaluation of adaptive filtering of MPEG system streams in IP networks". In: *IEEE International Conference on Multimedia and Expo 2000*, pp. 1313 - 1317, New York, 2000. ISBN: 0-7803-6536-4.

[5] N. W. Farrell *et al*., "Apod experiment 1 -final report. Technical Report Technical Memorandum 1311", BBN Technologies LLC, May 2002.

[6] Partha P. Pal, Franklin Webber *et al*., "Defense-Enabling Using Advanced Middleware: An Example". *Proceedings Milcom 2001*, pp. 92-102, October 28-31, Tysons Corner, Virginia, 2001. ISBN:0-7803-7225-5.

[7] Mungee S, Surendran N, Krishnamurthy Y, Schmidt DC, "The design and performance of a CORBA Audio/Video Streaming Service", in *IEEE Proceedings of the Hawaiian International Conference is System Science*, pp. 8043-8059, January 2001.

[8] E. Lamboray, A. Zollinger *et al*., "Interactive multimedia streams in distributed application", *Computer & Graphics 27*, pp. 735-745, 2003.

[9] Yu T-P, Wu D., Meyer-Patel K., Rowe L., "dc: A live webcast control system". In: *Proceeding IS & T/SPIE Symposium on Electronic Imaging: Science & Technology, Multimedia Computing and Networking, vol. 4312*, pp. 111-122, San Jose, CA, January 2001.

[10] Y. Krishnamurthy, Kachroo *et al*., "Integration of QoS-enabled Distributed Object Computing Middleware for Developing Nextgeneration Distributed Applications", *Proceedings of the ACM SIGPLAN Workshop on Optimization of Middleware and Distributed Systems (OM 2001)*, , pp. 230-237, Snowbird, Utah, 2001. ISBN: 1-58113-425-8.

[11] EKarr D.A., Rodrigues C. *et al*., "Application of the QuO quality-of-service framework to a distributed video application" In: *Proceedings of the Third International Symposium on Distributed Objects and Applications*, pp. 299-309, Rome (Italy), September 2001

[12] F. Garcia-Sanchez, A. J. Garcia-Sanchez *et al*., "Use of COTS for Raw Video Integration", *2003 IEEE Pacific Rim Conference on Communications Computer and Signal Processing*. pp. 671-674, Victoria (Canada), August 2003. ISBN: 0-7803-7978-0.

Diseño y validación de una herramienta de carga distribuida para servicios de audio y vídeo en Internet

M. Vilas, V. García, D. Melendi, X. G. Pañeda, R. García, I. Rodriguez
Departamento de Informática. Universidad de Oviedo
Campus Universitario del Viesques. Sede Departamental Oeste
33204 – Gijón. Asturias. España.
Teléfono: 986 18 19 86 Fax: 986 81 21 16
E-mail: {vilasmanuel, victor, melendi, xabiel, garciaroberto, rodriguezisabel}@uniovi.es

Abstract. *All time availability and reliability are two of the most important aspects in present services on the Internet. It is extremely important to know the best configuration parameters and the behaviour of the service on stress periods. Furthermore, the design, evaluation and implementation of new service architectures or new technologies for this type of services on working environments are extremely complicated tasks. This paper presents the design and validation of a workload generator for audio/video services on the Internet. The workload generator can be configured to emulate a high number of simultaneous users, with different user behaviours and in a distributed fashion. We also present a test environment that, jointly with the workload generator, will allow service administrators or research teams to evaluate different aspects of these services, such as the quality perceived in the users side or the effects of different interconnection network configurations.*

1 Introducción

No han pasado muchos años desde que Internet era una red con fines meramente académicos, utilizada solamente por un grupo reducido de usuarios para intercambiar información relacionada con sus investigaciones y estudios. La aparición del World Wide Web y los navegadores gráficos ha provocado un enorme aumento tanto en el número de usuarios como en la cantidad de información intercambiada. Este incremento ha provocado el desarrollo de un gran interés comercial por parte de numerosas empresas intentando explotar la popularidad de la red de redes. El enorme incremento de ancho de banda que han sufrido las líneas de acceso de los usuarios en los últimos años ha permitido el desarrollo de nuevos servicios que tan solo hace unos años serían impensables. Entre estos nuevos servicios uno de los más destacados son los de audio/vídeo en Internet.

Las elevadas necesidades en cuanto a consumo de recursos, tanto en los equipos terminales (servidor y cliente) como en la red de comunicaciones y la obligación de mantener una calidad de servicio constante durante la duración de los accesos de los usuarios, hacen que la configuración óptima de este tipo de servicios sea una tarea complicada y crucial de cara a garantizar su correcto funcionamiento. Si a esto añadimos la elevada variabilidad que presenta el comportamiento de los usuarios reales, la demanda de los recursos no se mantiene constante, existiendo periodos de elevada demanda conocidos como periodos de estrés.

Actualmente, en la gran mayoría de los casos, la configuración de los servicios se realiza en base a su experiencia previa. Sin embargo, cada servicio tiene sus particularidades que provocan que los conocimientos adquiridos en un servicio no sean directamente extrapolables a otro. Además, este conocimiento adquirido se basa en el pasado, lo cual en un mundo tan cambiante como el de Internet, en el que la evolución tecnológica es constante y la cantidad de demanda crece día a día, puede provocar que los parámetros de configuración utilizados en el servicio no sean los más adecuados. En los servicios actuales, en los que la disponibilidad y fiabilidad son parámetros de importancia crucial, realizar cambios en un entorno real, al que los usuarios pueden estar accediendo, sin conocer previamente cómo va a responder el sistema, supone un riesgo potencial. Con el objetivo de aislar la prestación del servicio real de todos estos aspectos, debe recurrirse a la realización de pruebas sobre el funcionamiento del servicio en un entorno a escala, aislado y controlado. Dentro de este escenario debe incluirse una herramienta de generación de carga que emule los accesos de múltiples usuarios, de una forma distribuida y con diferentes patrones de comportamiento configurables. De esta forma, preguntas como ¿Cuál es el comportamiento del sistema bajo condiciones extremas de carga? ¿Cuál es el número máximo de usuarios con un comportamiento determinado que soporta nuestro acuerdo con el operador de red? ¿Cómo afectan a la calidad percibida por el usuario factores como los protocolos de encaminamiento o las técnicas de traducción de direcciones? ¿Cuál es la configuración más adecuada del servicio si nos planteamos una nueva arquitectura del mismo? pueden ser resueltas sin afectar al sistema productivo real.

En este trabajo presentamos una herramienta de generación de carga distribuida y configurable que, en conjunción con el escenario de *test* adecuado,

permite la evaluación del funcionamiento de servicios de audio y vídeo, tanto por parte de administradores de servicios reales (con el objetivo de minimizar el impacto sobre el servicio real de cualquier modificación sobre la arquitectura del mismo o del comportamiento de los usuarios) como por parte de grupos de investigación interesados en el tema (con el objetivo de mejorar aspectos tecnológicos del funcionamiento de dichos servicios).

El resto del artículo está organizado de la siguiente forma: en la sección siguiente se hará un recorrido por los trabajos anteriores relacionados con el tema. El diseño de la herramienta de carga se presentará en la sección 3. En la sección 4 presentaremos un escenario de carga flexible y completo sobre el que disponer la arquitectura de servicio deseada y realizar la generación de carga. La sección 5 se centrará en el desarrollo de un modelo analítico que nos permita validar el comportamiento de la herramienta de carga. En la sección 6 se planteará un caso de estudio y se analizaran los resultados obtenidos. Finalmente, las conclusiones y los trabajos futuros serán expuestos en las secciones 7 y 8.

2 Trabajos previos.

Uno de los aspectos básicos a estudiar de cara al diseño de una herramienta de carga flexible y configurable que reproduzca fielmente el comportamiento de poblaciones de usuarios reales es el estudio de los principales parámetros que caracterizan el comportamiento de dichos usuarios. Respecto a los estudios realizados sobre servicios de audio/vídeo *streaming* cabe destacar un primer grupo centrado en el análisis de características propias de servicios de tipo continuo como [1,2] basados en la experiencia anterior en el estudio de sistemas web clásicos. En un segundo grupo podemos incluir aquellos trabajos que tratan de caracterizar el comportamiento de dichos usuarios analizando parámetros específicos de servicios de audio/vídeo *streaming*. En [3] y [4] se analizan servicios de vídeo de ámbitos universitarios con objetivos educacionales. En estos trabajos se analizan parámetros como la popularidad de los vídeos (típicamente modelada con una distribución *Zipf*), la cantidad de audio/vídeo transmitida y el tiempo entre sesiones de un usuario. Pese al enorme interés del estudio realizado, los resultados obtenidos sobre el comportamiento de los usuarios están fuertemente influenciados por el tipo de carga ofertado. En [5] se realiza un análisis detallado de los patrones de comportamiento de los usuarios que acceden a un servicio comercial en directo. Aspectos destacables de este estudio son la gran cantidad de accesos con los que cuentan (más de 3,5 millones de accesos) y la conclusión de que existe una clara correlación entre los parámetros que definen los accesos al servicio y el contenido ofertado; las preferencias de los usuarios influyen enormemente en la forma de los accesos. En [6] se realiza el análisis del comportamiento de los usuarios de dos servidores de vídeo bajo demanda pertenecientes a una conocida empresa multinacional del sector de la informática. Además de analizar las peticiones de los usuarios en cuanto a su duración y las interacciones realizadas, este estudio presenta interesantes resultados respecto a la caracterización de la carga y la evolución de la popularidad, planteando métricas que permitan medir esta evolución. Como uno de los aspectos más destacados de este estudio podemos destacar el hecho de que en un servicio como el analizado, dada la naturaleza de navegación de los usuarios que acceden a él, la duración de las peticiones es independiente de la duración del vídeo. En [7] se presenta una herramienta de análisis y configuración para servicios de audio/vídeo *streaming*, que tiene la novedad de evaluar una serie de métricas novedosas y diseñadas de forma específica para los servicios de este tipo. Mediante el módulo de análisis de esta herramienta se ha estudiado el comportamiento de los usuarios de dos servicios reales, tv.lne.es y www.asturies.com gracias a la información obtenida de los logs de los servidores.

Una vez analizados los parámetros básicos que caracterizan el comportamiento de los usuarios y los recursos a los que acceden, el siguiente escalón es el diseño de la herramienta de carga propiamente dicha. Una parte de los generadores de carga realizados para sistemas de audio/vídeo en Internet, se basan en el modelado y simulación del sistema [8] no en la generación real de carga, centrando los datos de salida generados por la herramienta en el consumo de ancho de banda. En [9,10,11] se realiza una evaluación de las capacidades de servidores multimedia mediante la emulación software del comportamiento de usuarios reales. Finalmente, en [12], entre otros muchos aspectos relacionados con los servicios de vídeo bajo demanda, se define un modelo completo de un servicio de este tipo, presentando un enorme interés de cara al diseño de la herramienta de carga.

3 Diseño de la herramienta de carga

La herramienta presentada en este trabajo sigue el diseño conceptual planteado en la Fig. 1. Los objetivos de la herramienta de carga son tres: simular los accesos de un número elevado de usuarios, definiendo de una forma flexible su comportamiento y permitiendo la generación distribuida de carga. Con esta finalidad se ha divido la herramienta en cuatro módulos diferenciados: módulo de registro, módulo de coordinación, módulo de usuario y módulo de reproducción. Estos módulos pueden desplegarse en diferentes máquinas para su distribución a través de la red de comunicaciones, como puede verse en la Fig. 2.

El módulo de registro es el encargado de recibir las peticiones de registro de los módulos de coordinación interesados en participar en la generación de la carga. Su función es establecer un punto de comunicación que sea conocido por todos los módulos de

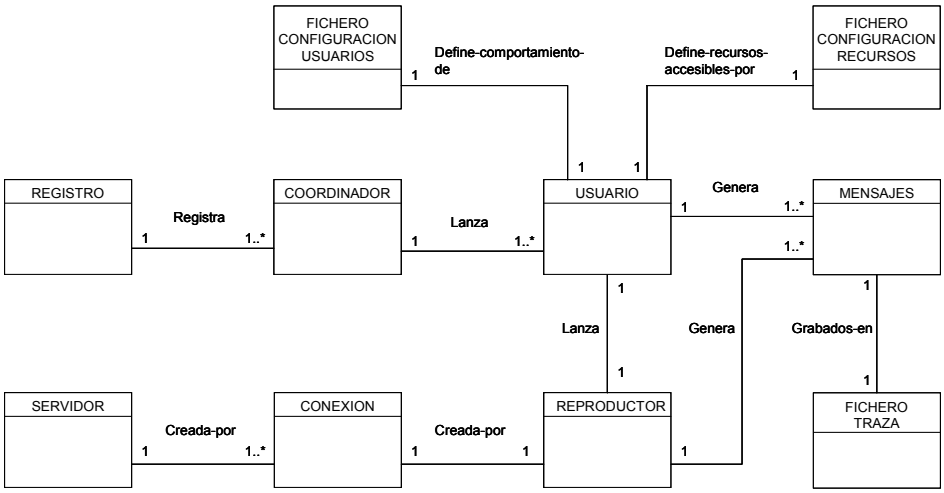


Figura 1: Modelo conceptual de la herramienta de carga

coordinación. Una vez que todos los módulos de coordinación se han registrado, uno de ellos se erige como iniciador de la carga, enviando un mensaje de inicialización al resto de los módulos de coordinación involucrados en la generación de la carga.

El módulo de coordinación es el encargado de establecer una referencia temporal común para la generación de la carga, posibilitando el análisis posterior de los resultados desde una referencia de tiempo común. En una carga puede existir un número variable de módulos coordinadores, los cuales, mediante un intercambio de mensajes inicial, establecen esta base de tiempo común. Tras esta tarea inicial, los módulos coordinadores se encargan de lanzar tantos módulos usuario como sea necesario, cada uno en un hilo diferente, independizando su funcionamiento. Una vez finalizada esta tarea, los módulos coordinadores se quedan a la espera de la finalización de los módulos de usuario.

El módulo de usuario es el encargado de emular el comportamiento de los usuarios. Dicho comportamiento se le indica mediante un fichero de configuración que analizaremos posteriormente.

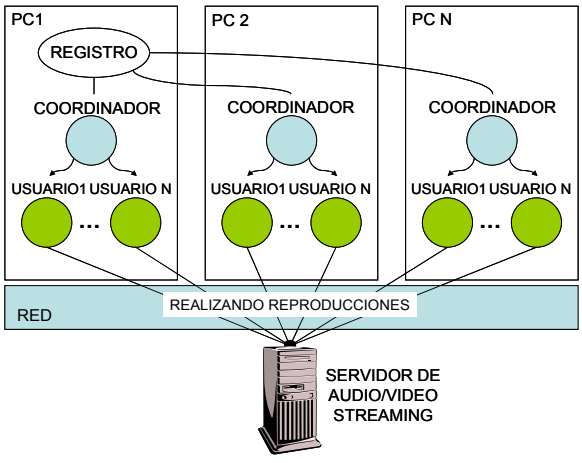


Figura 2: Estructura de la herramienta de carga.

Una vez que el módulo de usuario toma la decisión de lanzar una reproducción, mediante la generación de los valores pseudoaleatorios que marcan su comportamiento y que son obtenidos de un fichero de configuración, éste realiza una llamada al módulo de reproducción. Dicho módulo de reproducción está encargado de establecer la comunicación con el servidor y emular el intercambio de información típico entre un cliente y un servidor; solicita un recurso, realiza una negociación de sesión en la que establecen las condiciones de la reproducción y recibe el flujo de información. Asimismo, se encarga de avisar al servidor ante cualquier incidencia en la reproducción. El módulo desarrollado, basándose en la tecnología *HelixDNA* [13], realiza las mismas tareas que un reproductor real excepto el proceso de presentación al usuario final, el cual no es relevante para el objetivo del generador de carga y puede limitar en gran medida el número de usuarios simultáneos a emular. Una vez que el módulo de reproducción ha recibido el paquete del servidor, realiza el proceso de encolado del paquete, recoge la información necesaria y en el momento de presentarlo al usuario simplemente lo descarta.

3.1 Configuración de la carga

Con el objetivo de generar carga de una forma lo más realista posible, es necesario estudiar el comportamiento típico de un usuario. Los usuarios acceden al servicio durante sesiones compuestas de una o más reproducciones, con períodos de reflexión entre ellas como puede observarse en la Fig. 3. Cada una de las reproducciones comienza con una orden *play* y continúa hasta la finalización del vídeo o una interacción de *stop*. Durante las reproducciones los usuarios pueden realizar múltiples interacciones de *pause* y *play* con el servidor.

Para permitir la generación de carga de una forma flexible, se ha diseñado un documento XML que permite la definición de los parámetros estadísticos que marcan el comportamiento del usuario. En dichos documentos es posible definir las variables aleatorias

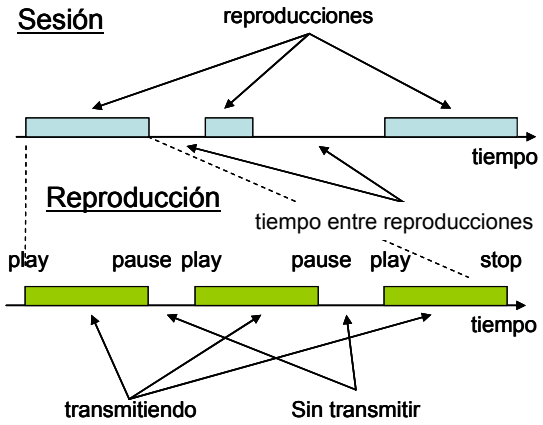


Figura 3: Comportamiento de los usuarios.

que marcan el tiempo entre peticiones de un usuario, la duración de dichas peticiones, el número de interacciones que realizará con el servidor, los instantes de tiempo en los que realizará dichas interacciones y la duración de las mismas,… como puede verse en el ejemplo planteado a continuación:

```
<user>
 <request>
  <distribution>
   <exp>3.0</exp>
  </distribution>
 </request>
 <session>
   <distribution>
    <exp>0.1</exp>
   </distribution>
 </session>
 <pause>
  <distribution>
   <discrete>
    <num_values>3</num_values>
    <value>0.9</value>
    <value>0.05</value>
    <value>0.05</value>
   </discrete>
  </distribution>
  <distribution>
   <pareto>
    <location>11.2</location>
    <shape>22.7</shape>
   </pareto>
  </distribution>
 </pause>
</user>
```

Este fichero definiría el comportamiento de usuario cuyo tiempo entre reproducciones se ajustase a una variable aleatoria exponencial, con tiempo de reproducción siguiendo una distribución también exponencial y cuyas interacciones con el servidor solo fuesen pausas, definidas en número por una variable aleatoria discreta y en duración por una distribución Pareto generalizada.

Otro aspecto a tener en cuenta en el comportamiento de los usuarios son los recursos que tienen a su disposición y la popularidad de los mismos. Esta información la obtienen los usuarios de un nuevo fichero, esta vez denominado de caracterización de recursos, cuyo objetivo es definir tanto la función Zipf que marca la popularidad de los mismos, como los URLs que definen los accesos. Nuevamente se ha diseñado un documento XML específico para la definición de estos parámetros. Un ejemplo, en el que la popularidad se define como una función Zipf de parámetro 0.667 aplicada a un conjunto de tres recursos diferentes tendría este formato:

```
<requested_urls>
 <zipf_like>
  <tita>0.667</tita>
  <num>3</num>
 </zipf_like>
 <url>rtsp://192.168.100.3/proyecto/video1.rm</url>
 <url>rtsp://192.168.100.3/proyecto/video2.rm</url>
 <url>rtsp://192.168.100.3/proyecto/video3.rm</url>
</requested_urls>
```

3.2 Datos de salida

Uno de los parámetros más importantes a la hora de evaluar el funcionamiento de un servicio de audio/vídeo en Internet es la percepción que tiene el usuario sobre el servicio. Con el objeto de analizar esta percepción, cada usuario emulado genera un fichero de traza en formato XML en el que se recogen los accesos realizados al servidor y las características de estos. En dichos ficheros se almacena información como el identificador del usuario dentro de la emulación, el recurso accedido, el instante de tiempo en el que se realizó dicho acceso y los principales eventos sucedidos durante la reproducción, como pueden ser inicio de un proceso de *buffering*, inicio de la reproducción, realización de una pausa o finalización de la reproducción. Asimismo, para cada reproducción se recoge información a nivel del intercambio de paquetes con el servidor, mostrando el número de paquetes total intercambiados, el número de paquetes recibidos en el intervalo adecuado, así como el número de paquetes retransmitidos y perdidos. También se muestra el ancho de banda medio consumido por la reproducción como puede observarse en el extracto de un fichero de traza típico:

```
<trace id_user="0">
 <playback num="0">
  <url time="1104915121">rtsp://192.168.100.3/video.rm</url>
    <begin>1104915122</begin>
    <playing>1104915122</playing>
    <buffering>1104915122</buffering>
    <playing>1104915125</playing>
    <stop>1104915132</stop>
    <statistics>
          <normal>277</normal>
          <recovered>8</recovered>
          <received>285</received>
          <lost>0</lost>
          <late>0</late>
          <bw>225000</bw>
    </statistics>
  <end>1104915132</end>
 </playback>
</trace>
```

4 Escenario de pruebas

Disponer de una herramienta de generación de carga flexible y configurable permite a los administradores de un servicio de audio/vídeo en Internet evaluar

ciertos parámetros de funcionamiento de su servicio y realizar una primera evaluación de nuevas arquitecturas. Sin embargo, sin conjuntar dicha herramienta con un escenario de pruebas completo y flexible, no podremos observar el efecto de todos los parámetros que tienen influencia sobre la calidad y disponibilidad del servicio en un entorno real. Con este objetivo se ha diseñado un escenario de pruebas que permite la evaluación de múltiples alternativas tanto en la red de interconexión como en la arquitectura del servicio.

4.1 Red de interconexión

La evaluación de arquitecturas avanzadas de servicio o la generación de carga distribuida imponen la necesidad de contar con una red de interconexión suficientemente compleja. Se han distribuido los equipos de comunicaciones alrededor de dos conmutadores LAN. Mediante la creación de VLANs (*Virtual Local Area Network*) y la asignación adecuada de puertos a estas VLANs se pueden recrear una enorme variedad de escenarios lógicos basados en un mismo escenario físico de interconexión. De esta forma, puede evaluarse el funcionamiento del servicio en entornos LAN, dentro de la red de un operador o desde la red de un operador diferente solamente modificando la configuración lógica de los equipos involucrados (direccionamiento IP, ficheros de configuración de los equipos de interconexión y asignación de puertos a VLANs en los conmutadores LAN).

4.2 Recogida y evaluación de los resultados

La información de log almacenada por los servidores multimedia incluye información de identificación de los usuarios, información sobre el consumo de recursos en la máquina e información acerca de la reproducción propiamente dicha. Toda esta información puede analizarse con la herramienta *Fesoria* [7], permitiendo analizar automáticamente la información almacenada en estos logs (desde decenas de entradas a varios miles por cada carga realizada) y obtener conclusiones sobre la calidad de servicio percibida. Además, dicha herramienta evalúa toda una serie de métricas diseñadas específicamente para el análisis de servicios multimedia.

Acerca del comportamiento de la red de interconexión, mediante la utilización del protocolo SNMP, es posible obtener de forma automatizada y almacenar en ficheros de log información como pérdidas de paquetes, ancho de banda consumido, carga de CPU,… en distintos puntos de la misma.

Respecto a la percepción del servicio por parte de los usuarios, ésta se puede analizar a partir de los ficheros de traza generados por los usuarios, y cuyo formato XML ya hemos mencionado en la sección de diseño de la herramienta. Se ha diseñado una plantilla XSL que genera gráficos SVG a partir de dichos ficheros de traza, permitiendo el análisis visual e intuitivo de entre decenas y varios miles de ficheros que de otra manera sería muy dificultoso. La Fig. 4 muestra un ejemplo de los gráficos generados donde se exponen las estadísticas a nivel de paquetes intercambiados para cada usuario y para cada una de la reproducciones realizadas por el usuario.

5 Validación de la herramienta

De cara a validar los resultados obtenidos del generador de carga, una opción es recurrir a la teoría de colas y la resolución analítica de modelos.

Obtener un modelo resoluble analíticamente y genérico de la herramienta de carga es una tarea extremadamente dificultosa. Sin embargo, si podemos suponer que el ancho de banda consumido por la carga generada se encuentra lejos de la capacidad máxima de las líneas de comunicaciones y que el servidor tiene suficiente capacidad para atender a la población de clientes bajo estudio, manteniéndose muy por debajo de su capacidad, con una red de comunicaciones basada en un solo *switch*, podemos obtener un modelo resoluble analíticamente. Bajo estas condiciones, podemos modelar el sistema como un sistema de colas con población finita [14], con número de recursos infinito y población igual al número de usuarios emulados por el conjunto de todos los equipos en los que se ejecuta el generador de carga.

En los modelos de población finita se trabaja con dos medidas; el tiempo de meditación y el tiempo de servicio en el sistema. Estas dos magnitudes son fácilmente asimilables a los tiempos entre accesos (tiempos de inactividad) y la duración de los mismos para los usuarios de un servicio de audio en directo. Si los usuarios tienen a su disposición solamente un recurso, el ancho de banda consumido por cada uno de ellos será igual al de todos los otros, es decir, el servicio se comportará uniformemente en todos los accesos.

Aplicando sobre nuestro modelo la fórmula de Little podemos relacionar el tamaño de la población, la tasa de generación de peticiones al servicio y las medias de los tiempos de reposo y actividad. Mediante este cálculo podemos obtener la tasa de generación de peticiones al servicio. De esta tasa y, de nuevo, recurriendo a la formula de Little, particularizada esta vez para el conjunto de infinitos recursos que modela nuestro servidor, podemos obtener el número medio de usuarios que están ocupando el servidor en un instante dado.

Una vez obtenido este valor sobre el número medio de usuarios, podemos calcular el consumo medio teórico de ancho de banda alcanzado en cada una de las pruebas con solo multiplicar el ancho de banda consumido por un solo acceso al servicio. Hemos de tener en cuenta la sobrecarga que introducen los protocolos de nivel inferior sobre la calidad con la

que está codificado el audio o vídeo; un archivo codificado a 20kbps consumirá una capacidad ligeramente superior debido a las sobrecargas introducidas por protocolos de nivel inferior.

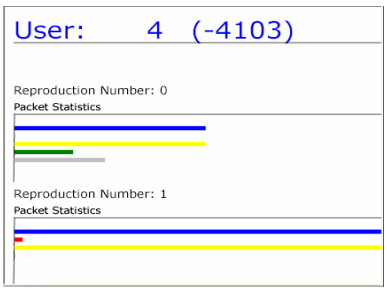


Figura 4: Gráfico SVG con estadísticas a nivel de paquetes intercambiados.

5.3 Resultados obtenidos

En la Fig. 5 podemos observar los resultados obtenidos, a partir del protocolo SNMP en las interfaces del equipo de conmutación LAN, durante una emulación de carga de 2 horas, representados para 4 valores diferentes de población (50, 100, 150 y 200 usuarios). Los usuarios realizaban peticiones siguiendo una variable aleatoria exponencial de media 1200 segundos y el tiempo entre peticiones seguía una distribución exponencial de media 600 segundos. En la misma gráfica se han representado los valores teóricos calculados a partir del modelo, representados como líneas horizontales, obteniendo unos valores similares a los medidos en la carga real una vez superado el transitorio.

6 Caso de estudio

El periódico digital Asturies.com, disponible desde el dominio www.asturies.com, ofrece un servicio de radio a través de Internet con un éxito considerable; en un año de existencia contabiliza miles de accesos.

Instalando en el entorno de pruebas una réplica de este servicio podremos evaluar aspectos del funcionamiento del mismo que, de otra forma, trabajando sobre el servicio real, podrían provocar que los usuarios experimentasen problemas en sus accesos.

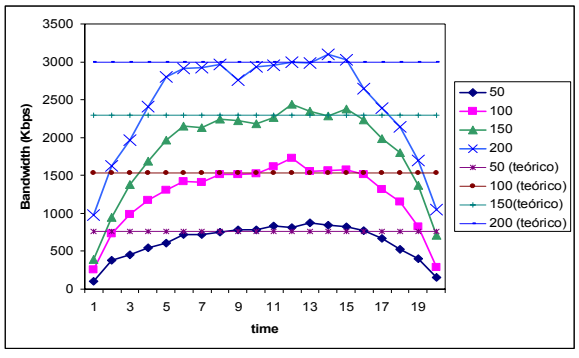


Figura 5: Comparación de los resultados teóricos y las mediciones del tráfico intercambiado.

6.1 Descripción del servicio

El servicio de radio de Asturies.com está basado en la distribución en vivo de contenidos almacenados previamente en el servidor. Asturies.com utiliza calidades de 11, 16 y 20Kbps para producir sus contenidos. En la misma máquina en la que se almacenan los contenidos se ejecuta un servidor de audio/vídeo en Internet y un slta (*Simulated Live Transfer Agent*). Este slta entrega el flujo en directo simulado al servidor, el cual se encarga de hacerlo disponible a los usuarios. El servidor instalado es el *Helix Universal Server*. No hay elementos de servicio adicionales a los mencionados y la entrega al usuario del flujo de información se realiza mediante tráfico *unicast*.

El servidor se encuentra en la red del operador Telecable, con el cual se ha establecido un acuerdo a nivel de servicio de 5Mbps garantizados.

Una de las características más importantes sobre el comportamiento de los usuarios de este servicio, obtenida del análisis de los logs del servidor de audio/vídeo, es la duración de las peticiones realizadas por los usuarios. Estas peticiones siguen una distribución exponencial de media μ=1.200 segundos.

6.2 Descripción del caso de estudio

Dado que la licencia con la que cuenta el servidor permite cargas de hasta 10Mbps, el ancho de banda establecido en el acuerdo a nivel de servicio con el operador de comunicaciones es el factor limitativo sobre el máximo número de usuarios que el servicio puede atender. A través de mediciones obtenidas de la generación de carga en condiciones extremas comparadas con las actuales del servicio (10 minutos de tiempo entre peticiones de un mismo usuario), se ha obtenido que el número máximo de usuarios con dicho comportamiento que el servicio soportaría sería de 300.

Una forma de superar esta limitación es recurrir a técnicas *multicast* para transmitir la información y atender a los usuarios. De esta forma, se transmitiría un único flujo de información para todos los usuarios. A esto hay que añadir que el servidor utilizado, *Helix Universal Server*, soporta dos alternativas de cara a dar soporte *multicast* para el servicio. La más atractiva de entre ellas es la conocida como *Back Channel Multicast* que aporta sobre la opción *Scalable Multicast* el hecho de que el URI que utilizan los usuarios para acceder al recurso es el mismo que para el caso *unicast*, y que, en caso de que el cliente no tenga capacidad para recibir tráfico *multicast*, se le entregará el flujo vía *unicast*. En esta opción, los clientes mantienen un canal de control con el servidor, permitiendo el envió de información de control y estadísticas de reproducción sobre tráfico *unicast* desde cliente a servidor.

El objetivo de este caso de estudio es evaluar el ahorro en cuanto a ancho de banda de la opción *Back Channel Multicast* frente a la opción *unicast* pura, observando los efectos sobre la percepción del servicio en el lado del usuario y observando el consumo de los canales de control desde los clientes al servidor.

6.2 Resultados obtenidos.

En la Fig. 6 pueden observarse los resultados obtenidos para el tráfico medido a la salida del servidor utilizando tráfico *unicast* para cuatro valores de población diferentes; 50, 100, 150 y 200 usuarios.

En la Fig. 7 pueden observarse los resultados obtenidos para el ancho de banda consumido por las transmisiones *multicast* con 4 valores diferentes de población; 50, 100, 150 y 200 usuarios. Puede observarse que el consumo de ancho de banda es constante para todos los valores de población y muy inferior al caso de transmisión *unicast*.

En la Fig. 8 puede observarse el tráfico consumido por los mensajes *unicast* de los clientes al servidor durante las emulaciones de carga para tráfico *unicast* (líneas discontinuas) y *multicast* (líneas continuas). Podemos destacar que el ancho de banda consumido por estas transmisiones *unicast* de los usuarios hacia el servidor es siempre menor o igual que la obtenida en el caso del envío de datos de servidor a cliente sobre flujos *unicast*.

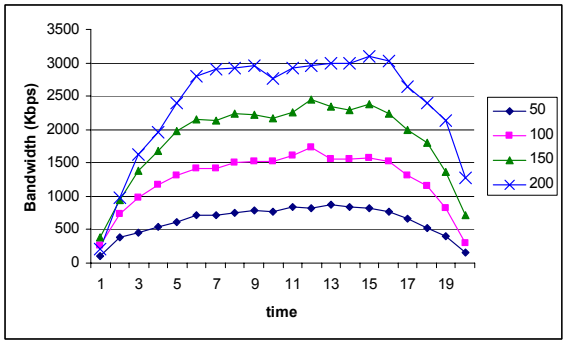


Figura 6: Tráfico enviado desde el servidor a los clientes utilizando transmisión *unicast*.

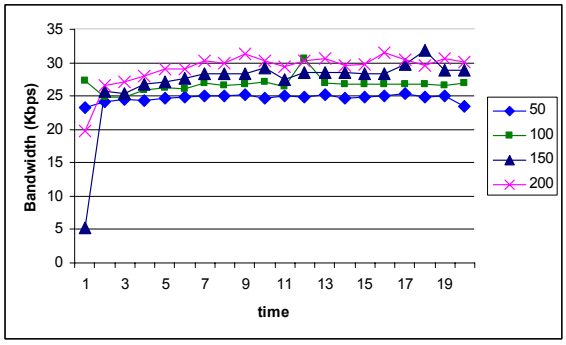


Figura 7: Tráfico enviado desde el servidor a los clientes utilizando transmisión *multicast*.

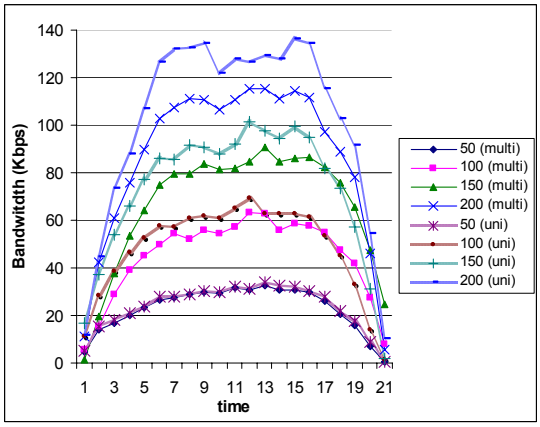


Figura 8: Tráfico enviado desde los clientes al servidor. (transmisión *multicast* y *unicast*)

Para el caso de transmisión *multicast* puede observarse cómo para valores de carga superiores a 50 usuarios, el ancho de banda consumido por los mensajes *unicast* de clientes a servidor (canales de control, sincronismo y recuperación de errores) supera el ancho de banda consumido por el por propio flujo de información *multicast*.

7 Conclusiones

El despliegue y configuración de servicios de audio/vídeo en Internet es un proceso complejo debido al elevado consumo de recursos necesario y la gran variabilidad existente entre el comportamiento de los usuarios de diferentes servicios. Si a esto añadimos que la disponibilidad y fiabilidad del servicio son parámetros de importancia vital para el éxito del mismo, disponer de una herramienta de generación flexible de carga puede convertirse en un aspecto que añada valor al servicio prestado. Disponer de un escenario de *test* en el que los gestores de dichos sistemas y los administradores de las redes de comunicaciones evalúen futuras soluciones antes de pasarlas a los entornos reales, sin afectar lo más mínimo al usuario, puede ser un factor vital en el éxito del servicio en un entorno tan cambiante y en el que la tecnología evoluciona de una forma tan rápida como Internet.

La herramienta presentada en este trabajo cumple todas esas características, permitiendo una experiencia satisfactoria para el usuario en sus accesos al servicio.

8. Trabajo futuro

En este trabajo se ha presentado el diseño y validación de una herramienta de carga. Pese a la validez y utilidad de los resultados alcanzados, hay un amplio espectro de temas pendientes de estudio.

La evaluación de arquitecturas complejas de servicio, en las que existe una red de servidores y *proxys* cuyo objetivo es acercar los contenidos a los usuarios y reducir el tráfico en ciertas zonas de la red de

comunicaciones es uno de ellos. Mediante la evaluación del funcionamiento de este tipo de arquitecturas queda abierta la posibilidad de analizar las consecuencias sobre los costes y los acuerdos establecidos entre la redes de diferentes operadores.

Otro aspecto en el que se puede profundizar es la parametrización del comportamiento de usuarios reales. Mediante el análisis de los accesos podemos extraer un patrón de comportamiento lo más cercano posible a la realidad, abriendo unas enormes expectativas en cuanto a la evaluación basada en predicciones del funcionamiento de servicios reales. Esta fase se contempla en algunos estudios previos como un aspecto de importancia significativa en un procedimiento completo de gestión de un servicio de audio/vídeo en Internet.

Respecto a la evaluación del servicio con tráfico *multicast*, un caso típico en la redes actuales, sería que solo una parte de los usuarios dispusiesen de capacidad *multicast*. La evaluación del funcionamiento del sistema en entornos mixtos queda abierta como trabajo futuro.

Finalmente, respecto al proceso de validación, sería de gran interés el desarrollo e implantación de un plan de pruebas completo y extenso.

Agradecimientos

Esta investigación ha sido financiada por el operador de comunicaciones Telecable de Asturias S.A.U y por el periódico La Nueva España dentro de los proyectos NuevaMedia, Telemedia, ModelMedia y MediaXXI y el Programa Nacional De Investigación dentro del proyecto INTEGRAMEDIA (TSI2004-00979).

Referencias

[1] C. Griwodz, M. Bär, Lars C. Wolf. Long-term Movie Popularity in Video-on-Demand System. ACM Multimedia. Seattle. 1997.

[2] D. Loguinov, H. Radha. Measurement Study of Low-bitrate Internet Video Streaming. ACM SIGCOMM Internet Measurement Workshop. 2001.

[3] M. Chesire, A. Wolman, G. Voelker, H. Lavy. Measurement and Analysis of a Streaming Media Workload. USENIX Symposium on Internet Technologies and Systems. 2001.

[4] Jussara M. Almeida, Jeffrey Krueger, Derek L. Eager, Mary K. Vernon. Analysis of Educational Media Server Workloads. NOSSDAV 2001. Port Jefferson, NY, EEUU. 2001.

[5] Eveline Veloso, Virgilio Almeida, Wagner Meira, Azer Bestavros, Shudong Jin. A Hierarchical Characterization of a Live Streaming Media Workload. Proceedings of the SIGCOMM Internet Measurement Workshop, Marseille, France, Nov 2002.

[6] Ludmila Chekasova, Minaxi Gupta: Analysis of Enterprise Media Server Workload: Access Patterns, Locality, Cotent Evolution and Rates of Change, IEEE/ACM Transactions on Networking, 2004.

[7] Xabiel Garcia Pañeda, David Melendi, Manuel Vilas, Isabel Rodríguez, Ricardo Bonis. Fesoria: An integrated tool for performance and content analysis, SLA evaluation, management and smart presentation for video-on-demand services. International Conference on E-Business and Telecommunication Networks (ICETE). Setubal, Portugal 2004.

[8] Shudong Jin and Azer Bestavros. GISMO: Generator of Streaming Media Objects and Workloads. ACM SIGMETRICS Performance Evaluation Review. 2001

[9] Mudashiru Busari, Carey Williamsom. ProWGen: A Synthetic Workload Generation Tool for Simulation Evaluation of Web Proxy Caches. Computer Networks: The International Journal of Computer and Telecommunications Networking. Pages 779-794. Volume 38, Issue 6. 2002.

[10] José Arias, Francisco Suarez, Daniel García, Xabiel G. Pañeda, Victor García. Evaluation of Video Server Capacity with Regard to Quality of the Service in Interactive News-on-Demand Systems. Protocols and Systems for Interactive Distributed Multimedia (IDMS). Lecture Notes in Computer Science, LNCS2515, Springer Verlag. Coimbra, Portugal, 2002.

[11] José Arias, Francisco Suárez, Daniel García, Xabiel G. Pañeda, Manuel García, Victor García. Evaluación de la capacidad del servidor de vídeo en función de la calidad del servicio en sistemas de noticias bajo demanda interactivos. CITA. Mérida, Venezuela. 2002.

[12] Javier G. Pañeda. Estudio de la Tecnología de Streaming como base para la Evaluación de Servicios de A/V Bajo Demanda Interactivos. Trabajo de investigación de los estudios de doctorado. Director: Francisco Suárez Alonso. Departamento de Informática, Universidad de Oviedo. 2001.

[13] Helix Community: https://helixcommunity.org/

[14] Pazos Arias, J. J., Suárez González, A., Díaz Redondo, R. P. Teoría de Colas y Simulación de Eventos Discretos. Pearson Educación. 2003.

Implementación y Evaluación de la Redirección de Usuarios en CDN

Benjamin Molina, Carlos E. Palau, Manuel Esteve, Isidoro Alonso y Victor Ruiz
Departamento de Comunicaciones. Universidad Politécnica de Valencia
ETSI de Telecomunicación. C/ Camino de Vera S/N. Campus Universitario.
46022 – Valencia (Valencia)
Teléfono: 96 387 73 01 Fax: 96 387 73 09
E-mail: benmomo@doctor.upv.es

***Abstract**. Overlay networks are application-level networks aiming at fulfilling different services that require both application and network intervention. Content delivery networks (CDNs) are overlay networks that redirect network users to close nearby servers, often called surrogates, and offer therefore a reduced response time. This feature involves collecting network information, server status information and contet location in order to provide the optimal surrogate for a given arbitrary client accesing either web or streaming-based content. Current CDNs are propietary globally deployed systems that manage hundreds of surrogates, but their internal behaviour remain hidden , so it seems difficult to adequately evaluate the performance. This article describes a general implementation of a CDN focussed on the redirection performance analysis.*

1 Introducción

Históricamente los contenidos digitales se han mantenido en entornos de grandes servidores centralizados en una única ubicación geográfica. Este tipo de solución proporciona malas prestaciones en términos de escalabilidad, no proporcionando el mejor tiempo de respuesta para todos los clientes. Por lo tanto, se ha tendido a adoptar múltiples soluciones para la distribución escalable de contenidos: clusters [1], sistemas de web caching, redes de distribución de contenidos (CDNs) [2] y, más recientemente, estructuras del tipo P2P [3]. Sin embargo, las diferencias entre las arquitecturas de los anteriores sistemas son significativas.

Las CDNs son redes de nivel de aplicación sobre Internet o redes TCP/IP, diseñadas para mejorar dos métricas: tiempo de respuesta y throughput total del sistema [4]. La primera métrica es importante para los clientes y supone el primer parámetro de valoración de este tipo de sistemas, mientras que la segunda métrica supone la cantidad de peticiones que pueden ser satisfechas por unidad de tiempo. Los servidores delegados (surrogates) son elementos clave en el despliegue y funcionamiento de una CDN, se comportan básicamente como servidores de caché, pero con un gestor de contenidos centralizado controlando los contenidos y su ubicación. La política de gestión de una CDN determinará la cantidad de información almacenada en cada uno de los surrogates. Cuando un cliente realiza una petición de contenido a un servidor origen cuyo contenido está gestionado por una CDN, la petición se redirige al servidor óptimo para servir esa petición, a efectos de que el cliente experimente el menor tiempo de respuesta, al menos si este tiempo se compara con el experimentado al pedir el contenido al servidor origen.

El proceso de acceso a contenidos se puede dividir en dos subprocesos: distribución de contenidos desde los surrogates y la redirección de las peticiones de los clientes al surrogate óptimo. Respecto a la distribución de los contenidos, las CDNs mejoran las prestaciones y la disponibilidad de los objetos multimedia a distribuir, acercándolos a los límites de la red y proporcionando servicios de réplica y de localización de contenidos.

El elemento clave en la mejora de las prestaciones proporcionadas por una CDN es la habilidad de dirigir las peticiones de cada cliente al surrogate más adecuado [5,6,7]. El algoritmo de redirección ha de considerar: (a) transparencia para los usuarios, (b) independencia del servicio, (c) independencia de la implementación, (d) escalabilidad, (e) flexibilidad en la política de selección del servidor e (f) independencia del proveedor del servicio. La reducción del tiempo de respuesta percibido por los clientes es el concepto básico de diseño en las CDNs.

No se ha prestado mucha atención a la implementación de las CDNs, y aunque algunos operadores y proveedores de servicio de CDNs como Akamai [8] o Speedera [9] ofrecen alguna documentacón describiendo su arquitectura de distribución de contenidos, se tiende a ocultar la implementación de forma propietaria como un elemento clave de éxito empresarial. Existen adicionalmente productos comerciales como ECDN (Cisco) y Content Director (Nortel Networks), así como desarrollos abiertos, como Globule [10] y CoDeeN [11].

Streaming CDN (SCDN) es una plataforma abierta desarrollada en la Universidad Politécnica de Valencia. El objetivo de este trabajo consiste en la distribución de objetos web de forma óptima así como de objetos multimedia en directo y bajo

demanda. SCDN presenta las siguientes características de implementación: (a) algoritmo de redirección de peticiones basado en información monitorizada, (b) desarrollo en JAVA con portabilidad entre plataformas, (c) escalabilidad para desplegar CDNs pequeñas, medianas o grandes, (d) utilización de filosofía COTS y (e) integración de un servidor de streaming multimedia para la distribución de vídeo y/o audio.

El resto del artículo se estructura como sigue: la sección 2 introduce la motivación y trabajo previo. Las secciones 3 y 4 describen la arquitectura y la implementación de la CDN, mientras que la sección 5 introduce los resultados del análisis. El artículo finaliza con las conclusiones y trabajo futuro.

2 Motivación y trabajo previo

La investigación previa del campo de las CDN se ha enfocado fundamentalmente en la evaluación de prestaciones de infraestructuras comerciales bien conocidas de CDNs. Los artículos de la literatura han investigado el uso y efectividad de las CDN, la reducción del tiempo de respuesta de forma empírica [7,12], efectividad de la redirección DNS [12,13], selección de servidores [6,7], o ubicación de los servidores [1]. Otros trabajos y contribuciones han tratado de modelar el comportamiento de las CDNs utilizando como parámetro la colocación de servidores [14] o la evaluación del tiempo de respuesta [7, 15-17].

Desde nuestro conocimiento no existen desarrollos abiertos de un sistema CDN. La aproximación realizada por Globule introduce una replicación de recursos orientada a objetos entre los diferentes componentes con el objetivo de crear una red de nivel de aplicación que los autores han denominado CDN centrada en usuario.

En cuanto al streaming y las CDN hay diferentes trabajos relacionados. PRISM proporciona identificación, gestión y descubrimiento de contenidos así como mecanismos de redirección para soportar streaming de alta calidad sobre una CDN basada en IP [18]. TVCDN es otro trabajo relacionado con streaming y CDN, aunque de momento no es más que una declaración de intenciones en el que los autores proponen un sistema de gestión de contenidos para la distribución de TV; no se desarrolla una infraestructura de CDN, sino que se supone que la misma se encuentra disponible. Otro sistema es MARCONINet [19], que proporciona una infraestructura para la distribución de audio a usuarios fijos y móviles utilizando proxies multimedia y gestión de contenidos, pero sin utilizar una CDN propiamente dicha. Y, finalmente en [20], los autores presentan diferentes técnicas y procedimientos para desarrollar una CDN orientada a la distribución de streaming a usuarios móviles.

3 Arquitectura de referencia

Una descripción de la arquitectura genérica de una CDN se puede localizar en diferentes contribuciones de la literatura [21]. La arquitectura utilizada en el sistema desarrollado en la UPV se representa en la Fig. 1. La arquitectura se basa en la utilización del protocolo HTTP y de los protocolos de streaming del IETF. La aplicación de la CDN se dirige a la distribución de objetos web y streaming de objetos multimedia. Los principales componentes de la arquitectura son: servidores origen, surrogates, clientes, red de distribución desde el origen, red de distribución a los clientes, gestor de contenidos y redirector.

Los servidores origen pertenecen a los propietarios/distribuidores de los contenidos, y contienen la información a ser distribuida o accedida por los clientes. Esta información se puede clasificar utilizando diferentes criterios. El criterio utilizado en este trabajo distingue entre contenido estático y dinámico. Los surrogates son réplicas totales o parciales de los servidores origen, actúan como servidores de caché con la habilidad de almacenar y distribuir contenidos. Los clientes son usuarios individuales con PC o con dispositivos especiales que solicitan y descargan contenidos almacenados en algún lugar de la CDN. Las CDNs habitualmente suelen tratar con clusters de clientes más que con clientes individuales. Los clusters suelen experimentar una latencia similar, así como las mismas restricciones de ancho de banda. La principal tarea del gestor de contenidos es la de controlar los objetos almacenados en cada uno de los surrogates, proporcionando esta información al módulo redirector para que cada cliente cuando emita una petición sea servido por el surrogate óptimo. El módulo Redirector proporciona inteligencia al sistema, porque se encarga de estimar el surrogate más adecuado para cada petición y cada cliente. Se descompone en tres módulos diferenciados: (a) CDN_{DNS}, asociado a la redirección (b) Monitor, correspondiente a la monitorización y (c) Algoritmo de Redirección, que se comentará posteriormente.

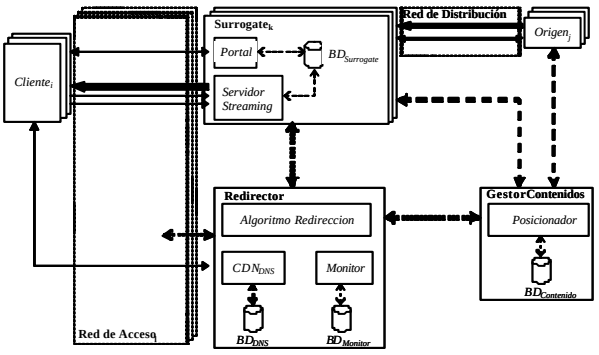


Figura 1. Arquitectura general de CDN.

4 Implementación de la CDN

La implementación de la CDN realizada ha consistido en la integración de diferentes elementos. En este artículo estamos interesados en la capacidad de redirección del sistema completo, por ello le hemos dedicado especial atención a la descripción del módulo DNS, el módulo Monitor, los agentes en los surrogates y principalmente al módulo del Algoritmo de Redirección, al ser elemento principal de la CDN.

4.1 Servidor DNS

El servidor DNS atiende y procesa todas las peticiones DNS relacionadas con los dominios gestionados por el sistema de la CDN. Cuando un cliente accede a la CDN, el primer paso consiste en proporcionar al cliente la dirección IP del surrogate más adecuado para esa petición. Diferentes estrategias pueden ser empleadas, desde una asignación estática a otras más complejas teniendo en cuenta el estado de los servidores y de la red. La CDN implementada en este proyecto utiliza, al igual que las CDNs de diferentes operadores, la segunda de las alternativas.

4.2 Módulo Monitor

Este módulo tiene la tarea fundamental de recopilar y almacenar información relativa a las prestaciones del sistema. Este módulo utiliza diferentes procedimientos para realizar la tarea. Estima el RTT para prevenir la congestión de la red mediante utilidades COTS tipo ping, tracert o medidas de RTT procedentes de TCP. El módulo Monitor utiliza SNMP para medir: la utilización de la CPU, el consumo de memoria, el número de conexiones TCP abiertas en los surrogates, etc.

4.3 Agentes SNMP

Cada surrogate de la CDN ejecuta un agente que tiene como tarea la recopilación de la información de estado local, así como el almacenamiento en su base de datos local. Si el sistema se comporta de forma correcta, una monitorización del tipo SNMP es suficiente, sin embargo el protocolo de gestión estándar proporciona poca información, con lo que los agentes desarrollados funcionan sobre SNMP pero de forma adaptada a nuestros requisitos.

4.4 Algoritmo de redirección

Dentro de un entorno CDN, un cliente debe ser redirigido al surrogate más apropiado; esto implica una cierta inteligencia que puede ser proporcionada por distintos mecanismos. Nuestra propuesta de CDN emplea redirección mediante DNS. El algoritmo de redirección, empleando la información capturada por el módulo Monitor, determina el contenido de la respuesta.

Revisando la traza de una sesión al acceder a una CDN, el cliente inicialmente contacta una página web donde puede seleccionar contenido adicional en formato streaming. Dado que dicha página inicial puede ser considerada de pequeño tamaño con pocas imágenes, resulta innecesario realizar tareas adicionales para determinar el surrogate más cercano; prácticamente cualquiera de ellos proporcionarán la página dentro de un reducido periodo de tiempo. Por ello, en esta situación sólo resulta necesario redirigir a un cliente a un surrogate poco cargado. Por el contrario, cuando se solicita un contenido de tipo streaming, resulta necesario un surrogate cercano para minimizar el efecto de la impredictibilidad de la red. Por tanto, existen dos modos de funcionamiento del algoritmo de redirección dependiendo del tipo de contenido solicitado (web o streaming).

El primer modo evalúa periódicamente una función general que considera los recursos disponibles, como es la utilización de la CPU, el consumo de memoria y el número de conexiones. Las ecuaciones que configuran el algoritmo de redirección están basadas en [22] con ciertas diferencias y mejoras que se describirán en los sucesivos parágrafos.

Sea x(i,j) la carga del servidor i-ésimo en el intervalo j-ésimo:

$$x(i,j) = a_1 \cdot m(i,j) + a_2 \cdot \frac{l(i,j)}{L(i)} + a_3 \cdot \frac{c(i,j)}{C(i)} \qquad (1)$$

donde:

- m(i,j) representa la memoria consumida en el surrogate i-ésimo durante el intervalo j-ésimo;
- l(i,j) es la utilización de CPU del surrogate i-ésimo en el intervalo j-ésimo; L(i) es la máxima carga de CPU deseable en el surrogate i;
- c(i,j) representa el número de conexiones establecidas con el surrogate i en el j-ésimo intervalo de tiempo; C(i) es el número máximo de conexiones que se desea;
- a1, a2, a3 son factores en el intervalo [0..1] y a1+a2+a3 =1. Esto permite asignar diferentes pesos a la carga de CPU, a la memoria y a las conexiones dependiendo de su importancia.

La ecuación anterior devuelve un valor normalizado para cada uno de los N_s surrogates en cada intervalo temporal. Estos valores deben poderse ordenar para poder comparar dentro de cada intervalo la carga de cada uno de los servidores. Esta carga influirá decisivamente a la hora de asignar un orden de asignación de probabilidades de cara a encaminar una petición de un cliente a un surrogate.

Sea f(i,j) la fracción de nuevas sesiones que se redirigirán al surrogate i en el intervalo temporal j-ésimo:

$$f(i,j) = k_1[x(i,j-1)] \cdot f(i,j-1) \qquad (2)$$

donde k1 representa una función dependiente de x(i,j-1) y se escoge de tal modo que los servidores poco cargados obtienen un valor elevado (mayor de 1) mientras que los servidores cargados por encima de un umbral obtienen un valor reducido (menor de 1). De una forma más descriptiva, la función k1 debe contemplar las siguientes situaciones:

(a) fallo del servidor: si x(i,j)=0 esto implica un fallo en el servidor, por lo que no se deben encaminar peticiones a este surrogate, es decir, f(i,j) = 0,

(b) inicialización: en el proceso de inicialización o recuperación los servidores toman un valor inicial f_{min}

(c) congestión: se define una carga máxima que implica la no asignación de peticiones, f(i,j) = 0

(d) baja-carga: en condiciones de baja carga, la función k1 toma su máximo valor alrededor de 1.5; esto supone un incremento relativo del 50% en cada intervalo.

(e) estado permanente: en condiciones ideales se busca conseguir que un surrogate reciba la misma tasa de solicitudes de forma continua (k1 = 1).

La función k1 puede tener diferentes formas; en el caso de este artículo hemos optado por tomar una implementación lineal, que ha demostrado un comportamiento adecuado como se verá en el apartado de evaluación. De manera general, la aproximación linear se puede describir como:

$$k_1(x) = \begin{cases} k_{\max} & , x \le x_{\min} \\ \dfrac{k_{\max} - k_{mid}}{x_{mid} - x_{\min}} \cdot x & , x \in [x_{\min}, x_{mid}] \\ \dfrac{k_{mid}}{x_{\max} - x_{mid}} \cdot x & , x \in [x_{mid}, x_{\max}] \\ 0 & , x \ge x_{\max} \end{cases} \qquad (3)$$

En lo sucesivo se trabajará con los siguientes valores:

xmin= 0.03, kmax=1.5 xmid=0.5, kmid=1 xmax=0.9, kmin=0

El segundo modo de funcionamiento del algoritmo de redirección asociado a streaming también considera el estado de la red además de los recursos en los servidores. Este estado de la red se determina en términos relativos basados en la proximidad entre un cliente y los surrogates.

Sea y(i,j) la proximidad de red entre el i-ésimo surrogate y el j-ésimo cliente:

$$y(i,j) = b_1 \cdot \frac{d(i,j)}{D} + b_2 \cdot \frac{p(i,j)}{P} \qquad (4)$$

donde:

- d(i,j) representa el número de saltos entre el i-ésimo surrogate y el j-ésimo cliente;
- D es el número máximo de saltos deseado;
- p(i,j) representa la latencia de red (ping) entre el i-ésimo surrogate y el j-ésimo cliente;
- P es la latencia máxima deseada ;
- ß1, and ß2 son factores en el intervalo [0..1] con ß1+ß2=1. Esto permite asignar diferentes pesos.

Una vez más es necesario establecer un orden para realizar una asignación probabilística, mediante una función k2:

$$g(i) = k_2[y(i)] \qquad (5)$$

Un valor reducido de y(i,j) implica una mayor cercanía entre el surrogate i y el cliente j. La función k2 puede ser elegida de forma arbitraria, pero debe garantizar la condición anteriormente mencionada.

Una vez se han procesado las dos funciones anteriores, el segundo modo de operación del algoritmo de redirección las combina con los recursos disponibles en los servidores, obteniendo una expresión general:

$$h(i,j) = g_1 \cdot f(i,j) + g_2 \cdot g(i,j) \qquad (6)$$

siendo ?1 y ?2 factores en el intervalo [0..1] y ?1+?2 =1. Estos factores permiten asignar diferentes pesos. Nótese que en el primer modo de operación del algoritmo de redirección ?2=0, mientras que en el segundo modo de operación ambos factores toman valores superiores a cero.

5 Análisis de prestaciones

5.1 Maqueta de pruebas

La Fig. 2 muestra la arquitectura de la maqueta de pruebas, así como el papel que desempeñan cada uno de los elementos que la componen. Esta maqueta consta de tres subredes, cada una de ellas con un surrogate simulando un punto de presencia (PoP). Cada surrogate alberga contenido web y streaming, aunque por simplicidad en el análisis se asumirá que todo el contenido se encuentra replicado.

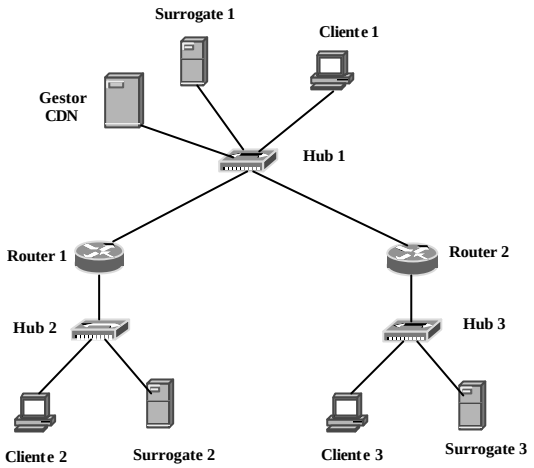


Figura 2. Maqueta de pruebas

En una de las subredes se encuentra el control de la CDN mediante un Gestor CDN que realiza, entre otras, tareas de redirección y gestión de contenidos.

Los clientes ejecutan un sencillo generador de carga que permite seleccionar diferentes patrones de solicitud, tanto de contenido web como streaming. La idea es que cada cliente simule en realidad un cluster de clientes, siendo capaces de causar un evento de tipo flash-crowd.

El análisis de prestaciones de la CDN se ha centrado en dos aspectos concretos: (a) estudio de la redirección DNS y (b) estudio de la redirección de contenido (streaming).

El primer test (redirección DNS) consiste en evaluar si las peticiones DNS son correctamente procesadas por el redirector encaminando cada petición al servidor más adecuado. Esto se consigue generando desde todos los clientes una tasa de peticiones constante y, al mismo tiempo, introducir una carga variable en cada surrogate, de tal manera que es posible comprobar el funcionamiento del algoritmo bajo diferentes escenarios de carga.

El segundo test (redirección de contenido) verifica el correcto comportamiento del modulo de redirección al solicitar peticiones relativas a contenidos. Se trata de un test similar al primero, pero en esta ocasión se solicita un contenido multimedia (streaming) con una carga variable en los surrogates; en este caso, ya se debe hacer una estimación del estado de la red.

5.2 Resultados

Como se ha comentado anteriormente, la maqueta consta de tres subredes (Ns=3), cada una de ellas con un surrogate y un cliente. Si este surrogate se encuentra descargado parece razonable que todas las peticiones provenientes de clientes en la misma subred sean redirigidas a dicho surrogate, y se encaminen a otro sólo cuando hay un incremento en la demanda (y por tanto de la carga). Este modo de proceder reduce la latencia percibida por los clientes y proporciona balanceo de carga entre todos los surrogates.

El algoritmo de redirección se implementa a través del servicio de DNS. Todas las medidas se han realizado para estudiar diferentes efectos sobre una base temporal de 10 segundos. Por otro lado, las gráficas sucesivas se han ordenado de izquierda a derecha en orden descendente.

El primer grupo de gráficas (Fig. 3) muestra la ejecución del algoritmo en ausencia de fallos en los servidores; además se ha introducido una carga de acceso a objetos web en cada cliente, variable e independiente. Esta variación afecta a la probabilidad de redirección de cada uno de los surrogates.

Para las medidas realizadas se ha tomado (empíricamente) un valor umbral $1/2 \cdot Ns$. Si el parámetro f es mayor de este valor la carga generada en la subred correspondiente será completamente asignada al surrogate local; en caso contrario, se realizará balanceo de carga entre todos los servidores.

El primer escenario (Fig. 3) evoluciona de la siguiente manera: al comienzo, la carga generada por los tres clusters de clientes es baja, por lo que cada solicitud se encamina al surrogate local. En el instante t=120 segundos, la carga genereada por el primer cliente aumenta, y de la misma forma lo hace el número de conexiones TCP pasivas y el uso de CPU en el primer surrogate. Esto modifica la ejecución del algoritmo y, una vez sobrepasa el umbral $f1$ (alrededor de 200 segundos), ciertas peticiones de clientes en la subred 1 se encaminan a las subredes 2 y 3, pues los surrogates de éstas aún permanecen descargados.

En el instante t=270 segundos, la carga generada por el cliente 2 aumenta, sin disminuir la carga producida por el cliente 1. Esto conlleva a un efecto similar en los parámetros correspondientes del segundo surrogate. Dado que el tercer surrogate es el menos cargado, éste ha de absorber no sólo peticiones del cliente 1, sino también del cliente 2 cuando $f2$ 'cae' por debajo de un umbral (alrededor de 400 segundos).

Finalmente, en el instante t=500 segundos, se aumenta la carga del tercer cliente, de forma que se establece un balanceo de carga similar al estado inicial.

Existen situaciones donde el acceso a un surrogate pueda parecer inalcanzable. La Fig. 4 representa el comportamiento del algoritmo ante la 'caida' de un surrogate, así como de su adecuada adaptación a esta situación. Al comienzo todos los surrogates se encuentran en un estado de carga similar, resultando en un valor parecido del parámetro f (alrededor de $1/N_s$). En el instante t=80 segundos, el cluster 1 deja de enviar solicitudes, por lo que el surrogate 1 deja de recibir peticiones. Esto se refleja en una variación

de los parámetros *f1*, *f2* y *f3*. En el instante t=600 segundos, el tercer surrogate sufre una caída. Este evento es detectado por el sistema de tal forma que las futures peticiones no se encaminarán a este surrogate hasta que no recupere su correcto estado de funcionamiento. Puede apreciarse que los dos primeros surrogates soportan una carga similar, dado que el parámetro *f* tiende a ecualizarse alrededor de 0.5.

Tras esta situación (instante t=1070 segundos), el segundo surrogate cae; esto se traduce en que todas conexiones serán soportadas por el surrogate restante, como puede apreciarse en las gráficas correspondientes.

La redirección de contenido también se ha testado cuando un cliente solicita un objeto multimedia (audio o video). El test se basa en la hipótesis que todos los surrogates disponen de dicho objeto. En caso contrario, el algoritmo de redirección tendrá que decidir entre servir el contenido desde un surrogate lejano que disponga de dicho contenido o desde un surrogate cercano (transfiriendo previamente dicho contenido). Esto dependerá del perfil con que se caracterice al cliente y se tratará como futuro trabajo. En cualquier caso, la premisa inicial (contenido replicado en todos los surrogates) no altera significativamente el comportamiento del algoritmo, más bien representa la adición de una característica adicional, sopesando unos factores frente a otros. Por otro lado, y para los objetivos del presente artículo, dicha asunción proporciona una mejor comprensión.

En el caso de la redirección por contenido se deben considerar parámetros adicionales (*y*, *g* y *h*) que se representan en la Fig. 5. Para cada cliente, el algoritmo debe crear una tabla independiente y-g-h para cada surrogate.

Dado que el contenido se encuentra replicado en todos los servidores, la solución más sencilla estriba en redirigir un cliente a su surrogate más cercano (misma subred) si éste no se encuentra excesivamente cargado. En este artículo, se ha establecido un umbral como se comentó anteriormente. De esta forma, un primer vistazo a la Fig. 5 revela que los parámetros *y*, *g* y *h* solo se calculan cuando el surrogate correspondiente (misma subred) se encuentra por debajo del umbral establecido. Para una comprensión más clara, la gráfica correspondiente al parámetro *h* se ha normalizado al valor máximo, por lo que el surrogate seleccionado siempre tendrá un valor igual a la unidad.

La Fig. 5 representa todas las gráficas obtenidas cuando un cliente ubicado en la segunda y tercera subred realizan peticiones de un contenido de tipo streaming. En el caso del cliente ubicado en la subred 3 (cliente 3), parece razonable redirigir su petición al surrogate 3, puesto que éste proporcionará la menor latencia. Esto está representado por el parámetro *g*; como se puede apreciar, la mayor parte del tiempo *y3* es inferior a *y1*, *y2*, por lo que el parámetro *g3* será mayor que *g1* y *g2*. Sin embargo, el parámetro final *h* también tiene en cuenta el valor de la carga en el tercer surrogate representado por *f3*, que se encuentra por debajo del umbral establecido en el intervalo [15-240]. Esto repercute en el hecho que el cliente 3 es redirigido en este intervalo temporal tanto al primer como al segundo surrogate. En el intante t=175 segundos, el segundo surrogate experimenta un carga aprecibale (el parámetro correspondiente decrece por debajo del umbral establecido), por lo que futuras peticiones procedentes del cliente 3 serán redirigidas al primer surrogate.

En el caso de un cliente ubicado en la segunda subred, la Fig. 5 representa un comportamiento similar del algoritmo de redirección. En esta situación, los surrogates 2 y 3 se encuentran cargados, por lo que el cliente 2 es inicialmente redirigido al primer surrogate. Nótese en la gráfica correspondiente del parámetro y que el cliente 2 percibe que la subred 3 se encuentra 'más cerca' que la subred 1, por lo que en cuanto el tercer surrogate pasa a estar menos cargado (instante t=120 segundos) todas las peticiones provenientes del Segundo surrogate pasan a encaminarse al tercer surrogate.

5 Conclusiones

La redirección de usuarios constituye una tarea crucial para que una red de distribución de contenidos pueda proporcionar un servicio escalable donde la latencia experimentada resulte mínima. En el caso de un sistema globalmente distribuido como una CDN, es necesario disponer de un conocimiento continuo de los servidores que la constituyen, así como del estado de la red, con el fin de redirigir a cada uno de los usuarios a un surrogate óptimo, ya sea porque se encuentren topológicamente cercanos o bien por tratarse de la mejor opción al proporcionar un menor retardo. El estudio del comportamiento del algoritmo que gestiona la decisión de la redirección de usuarios es fundamental para una correcta caracterización del servicio proporcionado. La gran mayoría de CDNs actuales son sistemas privados y cerrados, por lo que el estudio del rendimiento sólo puede realizarse por la interfaz externa que ofrece. En este artículo, se propone una implementación abierta con capacidad de actuación directa sobre el algoritmo de redirección, por lo que se accede también a la interfaz interna de la propia CDN. Asimismo, se han realizado una serie de pruebas reales sobre las que se ha testado el correcto funcionamiento del algoritmo.

Por otro lado, un sistema distribuido como una CDN dispone de un elevado número de componentes que constituyen su arquitectura y, por tanto, de un elevado número de factores parametrizables. La búsqueda de estos factores óptimos que mejoren el modelo ofrecido constituye el trabajo futuro a desarrollar.

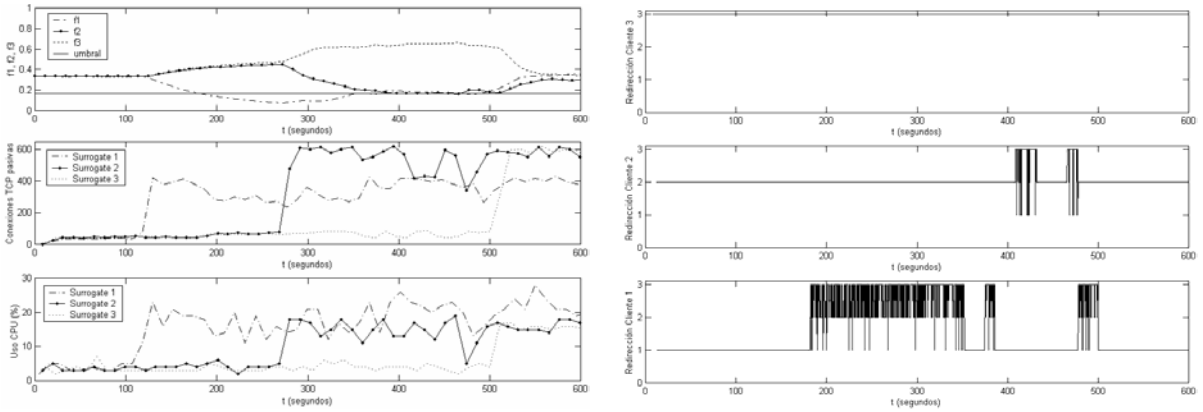


Figura 3. Escenario de trabajo en condiciones normales

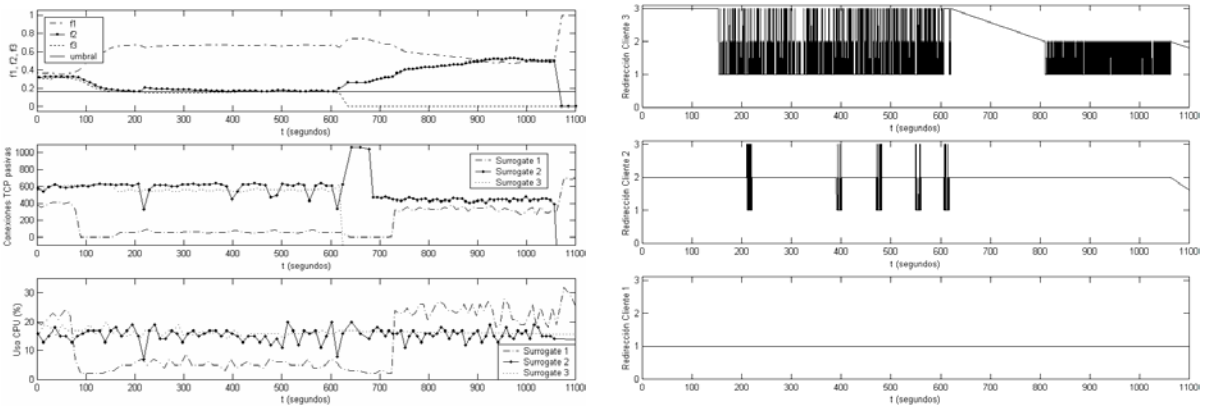


Figura 4. Escenario de trabajo con 'caidas' de surrogates

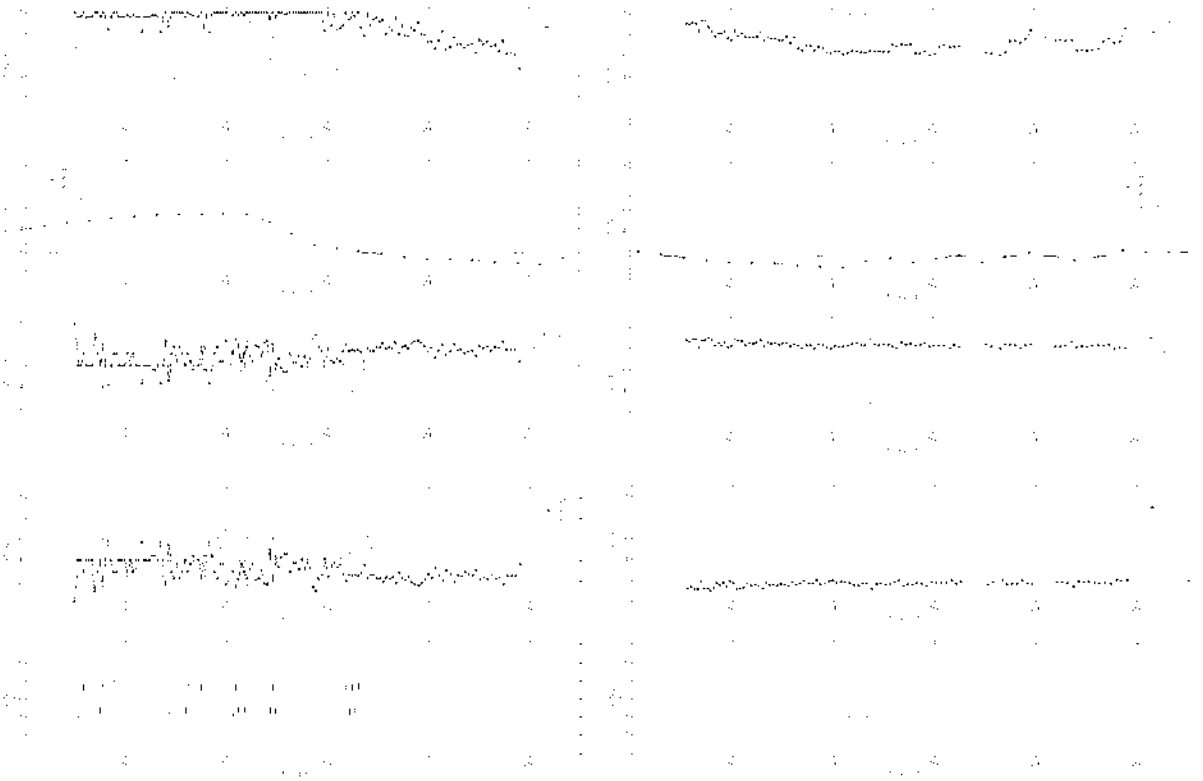


Figura 5. Escenario de trabajo con redirección de contenido streaming

Referencias

[1] P.S.M. Sayal, P. Vingralek, "Selection algorithms for replicated web servers". ACM SIGMETRICS Internet Server Performance Workshop. Madison (EEUU), Junio 1998.

[2] D. Verma. *Content Distribution Networks, an engineering approach*. John Wiley. ISBN:047122457X (2001).

[3] D. Liben-Nowell, H. Balakrishnan, D. Karger, "Analysis of the evolution of peer-to-peer systems". Proceedings of the twenty-first annual symposium on Principles of distributed computing, pp 233-242. Monterrey (EEUU), Julio 2002. ISBN:1-58113-485-1.

[4] S. Sariou, K.P. Gummadi, R. Dunn, S. Gribble, H. M. Levi. An analysis on Internet content delivery systems. Proceedings of the 5th symposium on Operating systems design and implementation, pp. 315-327. Boston (EEUU), Diciembre 2002. ISSN:0163-5980.

[5] K.L. Johnson, J.F. Carr, M.S. Day, M.F. Kaashoek. The measured performance of content distribution networks. 5th International Workshop on Web Caching and Content Distribution, Lisboa (Portugal), Junio 2000.

[6] R. P. Doyle, J. S. Chase, S. Gadde, A. M. Vahdat. The trickle-down effect: web caching and server request distribution. Proceedings of the Sixth International Web Content Caching and Distribution Workshop. Diciembre 2001. ISBN: 044450950X.

[7] J. Kangasharju, K.W. Ross, and J.W. Roberts, Performance Evaluation of Redirection Schemes in Content Distribution Networks, 5th International Workshop on Web Caching and Content Distribution. Lisboa (Portugal), Junio 2000.

[8] Akamai: http://www.akamai.com.

[9] Speedera Networks: http://www.speedera.com

[10] M. Szymaniak, G. Pierre, M. van Steen. Latency-driven replica placement. IEEE International Symposium on Applications and the Internet. Trento, Italia, Febrero 2005.

[11] K. Park, V. S. Pai, L. Peterson, Z. Wang, CoDNS: Improving DNS Performance and Reliability via Cooperative Lookups. Proceedings of the Sixth Symposium on Operating Systems Design and Implementation. OSDI '04. San Francisco, CA, Diciembre 2004.

[12] Z. Mao, C. Cranor, F. Douglis, M. Rabinovich, A precise and efficient evaluation of the proximity of web clients and their local DNS servers. Proceedings of the General Track: 2002 USENIX Annual Technical Conference, pp. 229–242. Monterrey CA (EEUU), Junio 2002. ISBN:1-880446-00-6

[13] B. Krishnamurthy, C. Wills, Y. Zhang. On the use and performance of Content Delivery Networks. Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement, pp. 169-182 . San Diego (EEUU), Agosto 2001. ISBN:1-58113-435-5

[14] C. Cameron, S. Low, D. Wei. High-Density model for server allocation and placement. Proceedings of the 2002 ACM SIGMETRICS international conference on Measurement and modeling of computer systems, pp. 152-159. Marina del Rey, CA (EEUU), Junio 2002. ISBN:1-58113-531-9

[15] D. Agrawal, J. Giles, D. Verma. On the performance of Content Distribution Networks. International Symposium on Performance Evaluation of Computer and Telecommunication Systems. Orlando (EEUU) , Julio 2001.

[16] M. Masa, E. Parravicini. Impact of Request Routing Algorithms on the Delivery Performance of Content Delivery Networks. Proceedings of the IEEE International Performance, Computing, and Communications Conference. Phoenix (EEUU), Abril 2003.

[17] B. Molina, C.E. Palau, M. Esteve. Modeling content delivery networks and their performance. Computer Communications 27(15), pp. 1401-1411, vol. 27, n. 15. 22 Septiembre 2004.

[18] A.Basso, et al. Prism, an IP-Based Architecture for Broadband Access to TV and Other Streaming Media. Procedings of the 10th Internationall Workshop Network and Operating System Support for Digital Audio and Video. Chapel Hill, Carolina del Norte (EEUU). Junio 2000.

[19] A. Dutta, H. Schulzrinne, Y. Yemini. MarconiNet: An architecture for Internet Radio and TV networks. Proc. International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV), pp. 241-245. New Jersey (EEUU), 23-25 Junio, 1999.

[20] S. Roy, M. Covell, J. Ankcorn, S. Wee, T. Yoshimura. A System Architecture for Managing Mobile Streaming Media Service. 23rd International Conference on Distributed Computing Systems Workshops (ICDCSW'03), pp. 408-413. Providencia, Rhode Island, (EEUU), 19-22 Mayo, 2003.

[21] G. Peng. CDN: Content Distribution Network. Technical Report TR-125. Experimental Computer Systems Lab, Department of Computer Science, State University of New York. Stony Brook, NY 2003

[22] M. Castro, M. Dwyer, M. Rumsewicz. Load balanacing and control for distributed World Wide Web Servers. Proceedings of the International Conference on Control Applications. Hawai (EEUU), 22-27 Agosto, 1999.

Una Propuesta de Arquitectura Adaptable para el Sistema de Encaminamiento de Peticiones de una CDN

Héctor Ossandón Díaz[1*], Encarna Pastor[**],
[*]Dpto. de Computación, Facultad de Ingeniería, Universidad de Tarapacá
18 de Septiembre Nº 2222. Arica-Chile
E-mail: ossandon@dit.upm.es
[**]Dpto. de Ingeniería de Sistemas Telemáticos, Universidad Politécnica de Madrid
ETSI Telecomunicación, Ciudad Universitaria s/n. 28040 - Madrid
E-mail: encarna@dit.upm.es

Abstract. *A Content Delivery Network (CDN) replicates the contents of origin Web servers to replica servers (surrogate servers) located in the edge of the network, close to the users. The Request Routing System of a CDN redirects user's requests to the best replica server in terms of response time. This redirection mechanism allows to improve performance metrics of the network such as access response time, Web server throughput and network bandwidth utilization. This paper presents a proposal of an adaptive architecture for the request routing system of CDNs that avoids additional network overhead. The proposed routing architecture includes measurement techniques from the user side. The chosen metric for the routing mechanism is based on latency from the point of view of the user. A simulation of the system is carried out using NS to evaluate system performances.*

1 Introducción y motivación.

La popularidad de Internet en los últimos años ha provocado que el tráfico de datos por la red aumente considerablemente, sin embargo, la actualización de su infraestructura tanto a nivel de enlaces como de nodos no sigue el mismo ritmo de crecimiento, lo que lleva a un resentimiento general de sus prestaciones que se traduce en tiempos de respuesta muy altos para ciertas aplicaciones [1].

Las Redes de Distribución de Contenidos (CDN, *Content Distribution Networks*) constituyen una perspectiva novedosa de cómo abordar el problema. Una CDN replica y distribuye parte o todos los contenidos de un sitio Web entre un conjunto de servidores réplicas (*surrogate servers*) dispersos geográficamente y ubicados en la periferia de la red, topológicamente cercanos a los sistemas de los usuarios finales [7]. El objetivo de una CDN es servir las peticiones del usuario desde aquel servidor réplica que proporcione el mejor servicio posible, de esta forma, se intenta reducir el tiempo de respuesta de las peticiones al evitar que los usuarios establezcan conexiones directas con los servidores Web origen.

La idea parece sencilla. Sin embargo, para que las CDN's se conviertan en una verdadera solución distribuida al problema de entrega de contenidos a través de Internet y contribuyan a mejorar la eficiencia y las prestaciones de ésta, en el diseño de su arquitectura se debe considerar una serie de aspectos críticos.

[1] Actualmente desarrollando la tesis doctoral en el Dpto. Ingeniería de Sistemas Telemáticos, ETSI de Telecomunicación, Universidad Politécnica de Madrid.

En este artículo se revisan los aspectos fundamentales para el diseño de la arquitectura de una CDN y se realiza una propuesta de arquitectura adaptable para su Sistema de Encaminamiento de Peticiones, así como su evaluación mediante simulación.

El artículo está organizado de la siguiente forma. En la sección 2 se da una visión general de la arquitectura de una CDN y se describen los aspectos más relevantes para el diseño de la arquitectura de su sistema de encaminamiento de peticiones. En la sección 3 se describe nuestra propuesta de arquitectura adaptable para el sistema de encaminamiento. En la sección 4 se presentan los resultados obtenidos de simulaciones hechas sobre una topología de red específica. La sección 5 presenta los comentarios finales del trabajo.

2 Aspectos de Diseño del Sistema de Encaminamiento de una CDN.

Las CDN's surgen como una evolución natural de los sistemas caché Web y se sustentan en toda la tecnología ya desarrollada para ellos y experimentada durante años de investigación, en [1] se puede encontrar una revisión de ellos. En términos generales, la arquitectura de una CDN está conformada por cuatro elementos básicos: una infraestructura de entrega de contenidos (servidores réplicas), un sistema de distribución, un sistema de encaminamiento de peticiones y un sistema de tarificación [3][4].

El sistema de distribución replica un conjunto selectivo de contenidos pertenecientes a uno o más servidores Web y lo almacena en servidores réplicas

que se encuentran más cercanos a los usuarios [2]. Cuando un usuario realiza una petición de contenidos, el sistema de encaminamiento redirige la petición al servidor réplica que pueda atenderla más eficientemente, que generalmente se refleja en aquel servidor réplica que le brinda el menor tiempo de respuesta. El sistema de encaminamiento sólo redirige las peticiones de los usuarios que corresponden a contenidos ya almacenados en los servidores réplicas por el sistema de distribución, para ello, debe existir algún tipo de realimentación entre ambos sistemas.

La infraestructura de entrega de contenidos interactúa con el usuario para entregar la copia del contenido que se solicita, además, interactúa con el sistema de tarificación para entregar información útil en el proceso de facturación del servicio. Finalmente, el sistema de tarificación se encarga de poner valor al servicio, los registros que éste elabora se pueden enviar al proveedor de contenidos (servidor origen) y/o al sistema de encaminamiento para utilizarlos en alguna política de encaminamiento. La Fig. 1, muestra las interacciones entre los elementos básicos de la arquitectura CDN.

2.1 Sistema de Encaminamiento.

En general, las prestaciones globales de una CDN están determinadas por su habilidad para redirigir las peticiones de los clientes al servidor réplica más apropiado en el momento, es decir, por la eficacia de su sistema de encaminamiento [5].

Un sistema de encaminamiento debe considerar en sus decisiones, no sólo información de tipo estática como localización geográfica o topología de la red, sino que también debe considerar información dinámica del estado de la red, como congestión de los enlaces y nivel de carga de trabajo en los servidores réplicas. El sistema de encaminamiento debe realizar dos funciones básicas: medir ciertos parámetros y, de acuerdo al resultado, encaminar las peticiones al servidor réplica que presente la mejor alternativa.

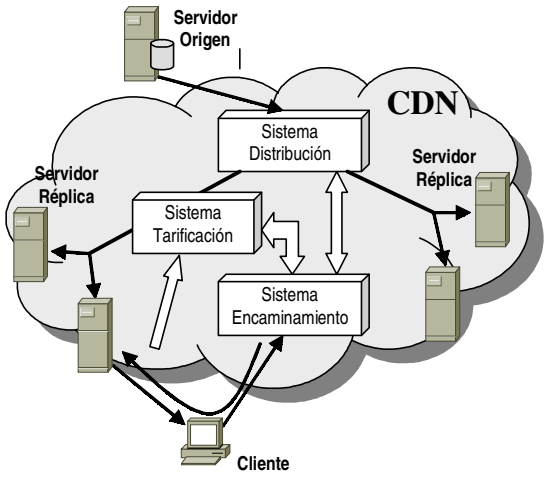


Figura 1: Arquitectura de una CDN.

2.2 Técnicas de Medición.

Entre las métricas más utilizadas para la elección del servidor réplica están: el número de saltos, ancho de banda, tasa de pérdida de paquetes, nivel de congestión, RTT (*Round-Trip-Time*), nivel de carga de trabajo y velocidad de procesamiento. Dependiendo del conjunto de métricas que se usen, el criterio de selección puede clasificarse en: estáticos, estadísticos y dinámicos [9].

Los criterios de selección estáticos se basan en la capacidad de los recursos hardware y de la topología de red, por lo que utilizan métricas que no varían durante un periodo largo de tiempo (semanas o meses). Este tipo de criterio no considera la demanda dinámica que sufren los recursos, por lo que son poco apropiados en un entorno dinámico como Internet. Sin embargo, la simpleza de su implantación y el poco overhead que producen los hacen bastantes atractivos, pero poco eficientes.

Los criterios de selección estadísticos utilizan la experiencia acumulada del comportamiento de una CDN, por lo que utilizan métricas tales como: latencia promedio y disponibilidad de ancho de banda promedio históricos. Para lograr obtener este tipo de métricas, generalmente, se utiliza la técnica de medición "pruebas pasivas" que consiste en examinar las cabeceras de los paquetes que fluyen a través de las conexiones TCP que se establecen, así como, las variables de estado de la pila del protocolo de cada conexión. Dependiendo de la variabilidad que presente la información, el nivel de confianza de este criterio de selección puede caer: a mayor variabilidad de los datos menor nivel de confianza.

Los criterios de selección dinámicos utilizan la técnica de medición de "pruebas activas". Ésta consiste en el envío de mensajes de prueba para medir en el momento las condiciones de las conexiones y servidores réplicas. Como he de esperarse este tipo de criterios produce mucho overhead, ya que por cada petición que se recibe, se deben enviar mensajes de prueba a cada uno de los servidores réplicas que guardan copias del contenido. Este tipo de criterio entrega una mayor confiabilidad, claro está, a un coste muy alto de overhead. Además, se debe tener presente que si las condiciones fluctúan muy rápido, la medición obtenida por el mensaje de prueba, puede que ya no sea válida para la transferencia de grandes archivos.

2.3 Esquemas de Medición.

Las pruebas o mediciones se realizan desde el servidor réplica hacia el usuario, "esquema de medición del lado servidor", o en sentido contrario "esquema de medición del lado cliente".

Los esquemas de medición del lado servidor son más utilizados debido a la facilidad de su implantación, sin embargo, presentan dos limitaciones

fundamentales. Primero, las mediciones hacia los usuarios que se encuentran asignados a un servidor en la periferia (DNS local, servidor caché, etc.) no reflejan fielmente las verdaderas prestaciones que percibe el usuario final. Y segundo, al ser los trayectos en Internet asimétricos, lo que realmente le interesa al usuario final son las prestaciones que otorga el trayecto de red en la dirección inversa, ya que en él se produce el mayor tráfico de datos.

Los esquemas de medición del lado cliente no tienen estas limitantes, ya que las métricas se obtienen desde el punto de vista del cliente, de manera que efectivamente reflejan las prestaciones que el cliente percibe. Sin embargo, este tipo de esquemas son muy difíciles de implantar eficientemente debido a la gran cantidad de overhead que producen [10].

2.4 Mecanismos de Encaminamiento.

Existen varios mecanismos de encaminamiento de peticiones. Según [6], se pueden clasificar bajo tres categorías: mecanismos basados en el Servidor de Nombres de Dominio (DNS), mecanismos a nivel de transporte y mecanismos a nivel de aplicación.

Los mecanismos de encaminamiento basados en el DNS utilizan un servidor DNS "especializado" que recibe las solicitudes de resolución de nombre del servidor DNS local al usuario y devuelve, de acuerdo a ciertas métricas y a la ubicación del usuario, la dirección IP del servidor réplica más próximo al usuario, en [11] se presenta un ejemplo. Una opción es que el servidor DNS especializado entregue, al servidor DNS local, varias direcciones IP correspondientes a servidores réplicas que pueden atender la solicitud y luego, el servidor DNS local las utilice cíclicamente (*Round-Robin*). Esta última alternativa busca distribuir la carga de trabajo por igual entre los servidores réplicas y a la vez aumentar la confiabilidad del sistema.

Aunque los mecanismos basados en servidor DNS son sencillos y no requieren grandes cambios en los protocolos actuales, presentan algunas limitaciones [6]: primero, debido a los múltiples niveles de redirección no hay buena escalabilidad; segundo, los TTL bajos provocan sobrecarga de solicitudes sobre los servidores DNS; y tercero, como las solicitudes al servidor DNS especializado se realizan a través de servidores DNS intermedios, la dirección del usuario es difícil de identificar.

Los mecanismos de encaminamiento de peticiones a nivel de transporte utilizan un router de peticiones para examinar la información disponible en el primer paquete de la petición del cliente y obtienen información como la dirección IP del cliente y el puerto TCP. Con esta información y otras métricas adhoc se determina el servidor réplica más apropiado para atender la petición del usuario. Generalmente este tipo de mecanismo se utiliza en combinación con alguno basado en servidor DNS para dar un nivel más fino de granularidad. Uno de los problemas de este esquema es que el router de peticiones hace de intermediario, al ser él quien establece la conexión con el servidor réplica, luego, todo el tráfico inicial pasa por él con el consiguiente retardo.

Los mecanismos de encaminamiento de peticiones a nivel de aplicación examinan los paquetes más exhaustivamente y a un nivel más alto. Con ellos se logra un control más fino en el encaminamiento de peticiones, por ejemplo a nivel de objetos. En [6] los clasifican en mecanismos de inspección de cabeceras y mecanismos de modificación de contenidos.

Los mecanismos de inspección de cabeceras pueden estar basados en la inspección del URL o de los identificadores específicos de sitio. Los mecanismos de modificación de contenidos re-escriben las referencias a objetos incrustados en los documentos Web, así, el cliente recupera los objetos desde el servidor réplica más apropiado. La reescritura de URL puede ser a priori o sobre demanda. Si es a priori, no se pueden tomar en cuenta las características específicas del cliente en la elección del servidor réplica, ya que la reescritura se realiza antes de conocer la identidad del cliente.

3 Arquitectura Propuesta.

Nuestra propuesta de arquitectura apunta a un esquema de medición del lado cliente. Esta decisión se basa en que este tipo de esquema refleja de mejor forma las prestaciones que percibe el usuario.

Las propuestas de arquitecturas para este tipo de esquema de medición en conocimiento de los autores, principalmente [8] y [9], utilizan pruebas dinámicas para obtener las métricas. Ya hemos mencionado el gran overhead en el tráfico que producen este tipo de pruebas, además, en una red con sobrecarga, la realización de pruebas dinámicas puede agravar la congestión no sólo en el tramo de red que lleva al servidor réplica que se elige, sino también en los tramos de red hacia los demás servidores réplicas que reciben las pruebas adicionales. Esto puede provocar que el resultado de las pruebas tarde mucho más tiempo en llegar al cliente, incrementando el tiempo que se invierte en la elección del servidor réplica.

Otro aspecto negativo de este tipo de pruebas, y que ya hemos mencionado antes, es que pierden parte de su efectividad cuando el archivo que se solicita es de gran tamaño y las condiciones de la red fluctúan muy rápidamente. Hemos querido volver sobre este aspecto ya que en el último tiempo ha adquirido mayor relevancia, debido al constante aumento en el tamaño de los archivos que se transmiten a través de la red (archivos con imágenes, de música, de películas, de presentaciones multimedia, etc.).

Para que una prueba en tiempo de ejecución sea efectiva, las condiciones de la red que ella midió deben mantenerse durante la transferencia del archivo

desde el servidor réplica hacia el usuario. En otro caso, si las condiciones de la red cambian, el servidor réplica que se eligió quizás ya no sea el más óptimo. Es claro que, en condiciones fluctuantes de la red, el impacto negativo de este aspecto va en directa proporción al tamaño del archivo que se transmite, es decir, la efectividad de las pruebas en tiempos de ejecución es menor si se transmiten archivos de gran tamaño en condiciones fluctuantes de la red que si se transmiten archivos de menor tamaño.

Considerando los puntos de los párrafos anteriores, sería interesante entonces tratar de obtener una predicción del comportamiento futuro inmediato de las condiciones de la red y de los servidores réplicas antes de tomar una decisión sobre la elección del servidor réplica. Los criterios de selección estadísticos pueden proporcionar la posibilidad de predecir el comportamiento futuro inmediato a partir de datos históricos. Nuestra propuesta considera entonces un esquema de medición del lado cliente con un criterio de selección de tipo estadístico. Para proporcionar una estimación más ajustada a la realidad, los datos estadísticos son acumulados por tramos horarios y clusters o zonas de usuarios. En Fig. 2 y Fig. 3 se muestran los componentes e interacciones de la estructura propuesta.

En términos generales, el usuario realiza una petición de contenidos al sistema de encaminamiento. Considerando la hora y la zona de la que proviene la petición, el sistema de encaminamiento revisa sus bases de datos locales de área y selecciona el servidor réplica que promete un mejor servicio (tiempo de respuesta menor). Luego, el usuario es redirigido al servidor réplica seleccionado para que recupere el contenido. Una vez que el usuario recupera el contenido, éste registra el tiempo de respuesta que obtuvo y envía el dato al sistema de encaminamiento para que actualice sus bases de datos locales.

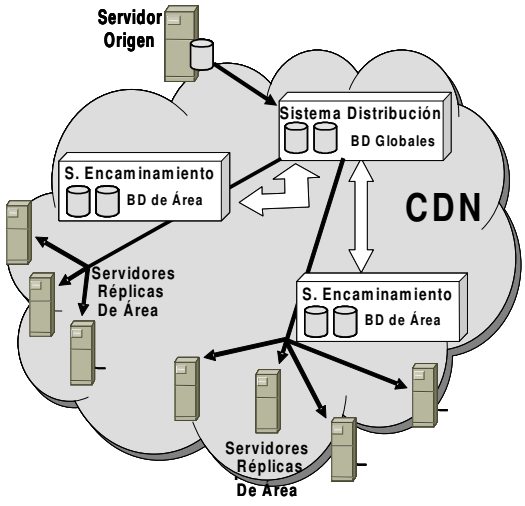


Figura 2: Arquitectura Adaptable, interacción Sistemas de Distribución y Sistema de Encaminamiento.

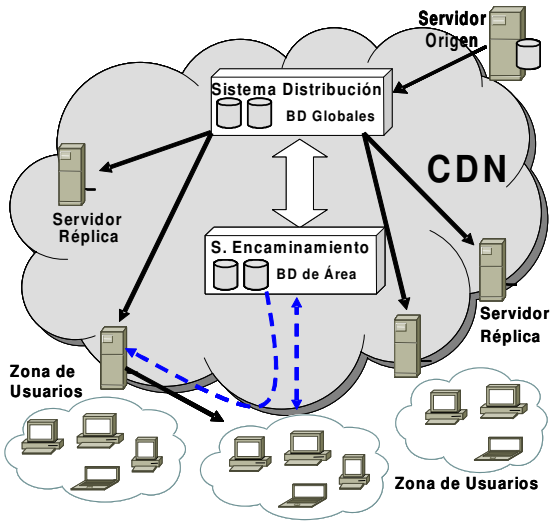


Figura 3: Interacción con zonas de usuarios de un área.

3.1 Métrica utilizada.

La métrica que se utiliza en nuestra propuesta para la elección del servidor réplica óptimo es el tiempo de respuesta final que percibe el usuario, ya que éste refleja en conjunto tanto el estado de la red como del servidor réplica.

El tiempo de respuesta final que percibe el usuario esta compuesto del tiempo que toma la consulta al DNS, del tiempo que demora el establecimiento de la conexión TCP/IP con el servidor réplica, del tiempo que transcurre entre el envió de la petición y la recepción del primer paquete de la respuesta (tiempo de latencia) y el tiempo que toma la recepción de los demás paquetes de la respuesta (tiempo de transferencia) [9]. Sin embargo, como los usuarios no perciben estos componentes de forma aislada, nuestra propuesta almacena los datos estadísticos de tiempos de respuestas finales obtenidos en cada una de las peticiones realizadas a los servidores réplicas.

Otra cuestión a tener presente es que los tiempos de respuesta que se obtienen de peticiones de archivos de distintos tamaños no son comparables directamente, por lo que dicha métrica se debe normalizar. Al igual que en [9] nuestra propuesta utiliza una regla de tres simple para normalizar el tiempo de respuesta; pero con la diferencia en que nosotros omitimos la consideración por separado de la componente tiempo de latencia y la incluimos como parte integral de la componente tiempo de transferencia. Lo anterior responde a dos cuestiones. La primera es que cuando se recuperan archivos grandes, el tiempo de latencia del primer paquete que se recupera no es significativo con respecto al tiempo de respuesta global. La segunda es que con ello, simplificamos tanto la medición del tiempo de respuesta global como el cálculo del tiempo de respuesta normalizado sin perder representabilidad de la métrica. Así, si consideramos que t_c es el tiempo de

conexión y t_t es el tiempo de transferencia que registra un usuario al recuperar un contenido de tamaño s desde un servidor réplica, el tiempo de respuesta normalizado $\hat{t}$ de recuperar un archivo de tamaño normalizado $\hat{s}$ será:

$$\hat{t} = t_c + t_t\left(\frac{\hat{s}}{s}\right) \qquad (1)$$

Esta operación se realiza en el cliente, ya que es éste quien mide t_c, t_t y s. Luego, el cliente envía al sistema de encaminamiento el resultado obtenido en (1), la franja horaria en que se realizó el servicio y la dirección IP del servidor réplica que lo atendió. Esta información se envía anexa en la siguiente petición del cliente para no generar una conexión TCP extra. Con estos datos el sistema de encaminamiento actualiza las bases de datos que maneja y que se describen en la siguiente sección.

3.2 Estructura de las Bases de Datos.

El sistema de encaminamiento de peticiones maneja bases de datos locales de áreas, que utiliza para predecir el comportamiento de los servidores réplicas durante las distintas franjas horarias y con respecto a cada una de las zonas de usuarios. El hecho de manejar bases de datos locales por área, permite que nuestra propuesta escale bien si el número de usuarios crece.

En términos prácticos, un área puede corresponder a un ISP (Proveedor de Servicio de Internet) y el sistema de encaminamiento de esa área podría estar al mismo nivel que sus DNS locales, incluso en la misma máquina física. Si el ISP posee muchos clientes o éstos se encuentran muy diseminados geográficamente, se pueden crear zonas de clientes utilizando los distintos mecanismos de agrupamiento de clientes que existen (*client clustering*) [12][13]. Así, se puede dar el caso de que existan áreas con una única zona de clientes (ISP pequeños y clientes concentrados geográficamente) o áreas con varias zonas de clientes (ISP grandes o con clientes diseminados geográficamente).

El objetivo de las bases de datos locales por área, es mantener, para cada servidor réplica, una predicción del tiempo de respuesta normalizado que se obtendrá en la siguiente petición desde una zona de usuarios. La Fig. 4 muestra la estructura de las tablas que componen una base de datos local de área.

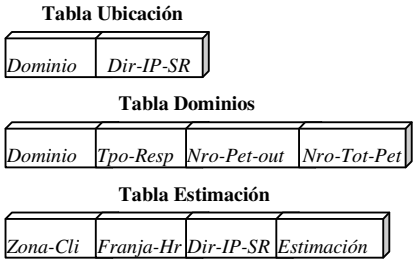


Figura 4: Estructura de la BD local de área.

Tabla Ubicación: Reúne información sobre qué servidores réplicas de área contienen copias de un dominio determinado. Esta tabla es alimentada por el Sistema de Distribución de la CDN mediante mensajes de actualización de la forma *(cod-op,dominio,dir-ip-sr)*, donde *cod-op* puede ser eliminar o insertar la entrada *(dominio,dir-ip-sr).*

Tabla Dominios: Reúne información sobre los dominios replicados. El campo *Tpo-Resp* refleja el tiempo de respuesta normalizado máximo en el que se debe atender una petición para este dominio, este dato puede ser consensuado con el proveedor de contenidos en un contrato de servicio o establecido de acuerdo a una estructura de precios. El campo *Nro-Pet-out* contabiliza el número de peticiones cuyos tiempos de respuesta normalizado han superado el valor del campo *Tpo-Resp* para un dominio durante una semana, éste lo utilizamos como un indicador de desempeño que guía el proceso de adaptación de la CDN. Finalmente el campo Nro-Tot-Pet refleja la cantidad total de peticiones recibidas durante una semana para ese dominio.

Tabla Estimación: Reúne la información que nos permite elegir al servidor réplica que promete un mejor tiempo de respuesta normalizado para una petición proveniente desde una zona de usuarios, en una franja horaria determinada.

3.3 Obtención de la Estimación.

Cuando el usuario solicita un contenido, su petición es capturada por el sistema de encaminamiento que examina la tabla de dominios y obtiene las direcciones IP de los servidores réplicas candidatos que pueden atender la petición. De acuerdo a la hora y día, se puede encasillar la petición dentro una franja horaria (*Franja-Hr).* Como se sugiere en [9] es recomendable calcular estimadores estadísticos para los diferentes tráficos de datos que se producen durante el día. Nuestra propuesta no establece la cantidad ni la duración de las franjas horarias, más bien se deja en libertad para que éstas se configuren en cada uno de los sistemas de encaminamiento de área, tomando en cuenta para ello los patrones de tráfico que generan las peticiones de los usuarios de las respectivas zonas en cada área. Un ejemplo puede ser un sistema de encaminamiento de dos franjas horarias: diurna (09-18hrs de lunes a viernes) y nocturna (en cualquier otro horario).

Con la dirección IP de los servidores réplicas candidatos, la franja horaria en que se produce la petición y la zona de usuarios de la que proviene la petición, el sistema de encaminamiento obtiene los tiempos de respuesta normalizados estimados para cada uno de los servidores réplicas candidatos y elige aquel que presenta la estimación de menor valor.

Como ya mencionamos en el apartado 3.2, una vez que el usuario es atendido por el servidor réplica, éste envía el valor $\hat{t}$ al sistema de encaminamiento que lo

utiliza para actualizar el campo *Estimación*. Nuestra propuesta plantea, basado en el comportamiento pasado de un servidor réplica con respecto a una zona de usuarios, predecir el tiempo de respuesta normalizado que se obtendrá en la próxima petición proveniente de la misma zona de usuarios y franja horaria, utilizando para ello un promedio exponencial de los tiempos de respuestas normalizados. Así, si T_n es el valor que se encuentra en el campo *Estimación* después que un servidor réplica ha satisfecho la enésima petición y $\hat{t}_n$ es el tiempo de respuesta normalizado enviado por el usuario tras esa enésima petición satisfecha, el nuevo valor del campo *Estimación* será:

$$T_{n+1} = (1-\alpha)\hat{t}_n + \alpha T_n \qquad (2)$$

con α entre 0 y 1. Si α se toma más cercano a 0, se le da más influencia al resultado de la última medición que al histórico de mediciones, si α se toma más cercano a 1 se le da más importancia al histórico de las mediciones que a la más reciente. En nuestro caso hemos optado por α=0.5, como medida intermedia, pero dependiendo de las condiciones de red, el valor se puede variar en cada área.

El promedio exponencial ya se ha utilizado para estimaciones de tráfico TCP, tanto en la estimación del RTT como en el ancho de banda [14] y ha resultado bastante aceptable, así su elección se ajusta a nuestro propósito y nos da cierta tranquilidad en la estimación.

3.4 Mecanismo de Adaptación.

El mecanismo de adaptación que planteamos es el siguiente. Un proceso que se ejecuta cada cierto intervalo de tiempo, por ejemplo cada semana, de preferencia en un horario de poca actividad, revisa la tabla *Dominios* para solicitar al sistema de distribución una posible adecuación de las copias de contenidos que se mantienen en los servidores réplicas del área respectiva.

Para solicitar al sistema de distribución nuevas copias de un contenido en su área, el proceso examina los campos *Nro-Pet-out* y *Nro-Tot-Pet* de la tabla de *Dominios*. Si la cantidad en *Nro-Pet-out* representa una porción de *Nro-Tot-Pe*t que se encuentra por encima de un primer valor umbral, el proceso solicita mediante un mensaje al sistema de distribución que proporcione una copia más del contenido del dominio respectivo, si se encuentra por sobre un segundo valor umbral solicita dos copias y así sucesivamente. Los valores umbrales que hemos definido son 10% para solicitar una copia, 20% para solicitar 2 copias, 30% para solicitar 3 copias, etc. Es claro, que al disponer de más servidores réplicas que puedan atender su petición, el usuario aumenta la probabilidad de encontrar un servidor réplica que no esté congestionado.

Para solicitar al sistema de distribución que elimine copias de un contenido en su área, el proceso examina sólo el campo *Nro-Tot-Pet*. Si el valor examinado es menor a un cierto valor umbral, el proceso solicita al sistema de distribución que elimine una réplica del contenido del dominio respectivo, siempre que no sea la única.

4 Simulación y evaluación.

En esta sección presentamos los resultados obtenidos en experimentos de simulación realizados sobre un modelo de topología de red específica, ver Fig. 5. Para construir los modelos de simulación utilizamos el simulador de redes NS-2 [15].

La zona de usuarios se modeló como varias instancias de la clase Http/Client conectadas a una instancia de la clase Http/Cache. Esta última se modificó para que cumpla funciones de servidor caché y sistema de encaminamiento de peticiones simultáneamente. En nuestros experimentos de simulación todos los contenidos se marcaron como no "cacheables" con el propósito de evaluar sólo el rendimiento del sistema de encaminamiento de peticiones. Sin embargo, con las modificaciones que se hicieron en el código, queda abierta la posibilidad de experimentar la influencia que pueden tener las distintas políticas de caching y mecanismos de consistencia en el rendimiento del sistema de encaminamiento de peticiones. De igual forma, poniendo el número de réplicas en cero, se anula la operación del sistema de encaminamiento de peticiones y se vuelve a operar sólo como servidor caché.

Los servidores Web y servidores réplicas se modelaron como instancias de la clase Http/Server. Para modelar la mayor distancia de los servidores Web en relación con los clientes, los enlaces finales que unen a los servidores Web con los routers se castigaron con tiempos de retardos altos, así éstos enlaces en realidad modelan varios saltos en la topología de la red. Esto es perfectamente válido como queda establecido en [16] y nos permite una mayor eficiencia en la ejecución de la simulación al reducir el tamaño de la red sin perder representatividad.

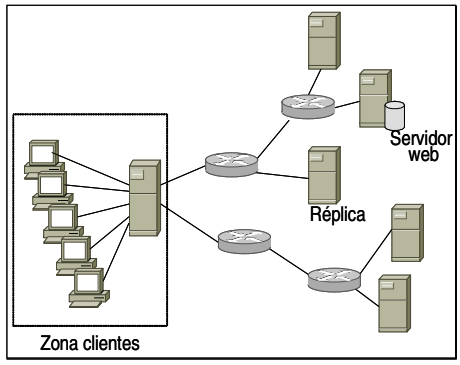


Figura 5: Topología de Red para simulación.

Para modelar la ejecución por parte de los clientes del cálculo del tiempo de respuesta normalizado, también se castigó a los enlaces que van desde éstos al servidor caché con un pequeño retardo.

La carga de trabajo para los experimentos de simulación se obtuvo de archivos de trazas disponibles en [17], estas trazas alimentaron directamente a los clientes del modelo. Para ello se utilizó la clase PagePool/ProxyTrace que permite conducir la simulación a través de trazas reales con un formato específico. Un problema que encontramos aquí fue que cuando utilizamos varios servidores Web para soportar la carga de trabajo global de la red, las peticiones a un mismo documento son dirigidas a diferentes servidores Web en forma aleatoria, es decir, no se mantiene la relación de pertenencia entre el documento que se solicita y el servidor Web que aloja la versión original. Así, hemos modificado esta clase para que esta relación de pertenencia se mantenga cuando se utilizan varios servidores Web. Luego, en nuestra implantación, un documento determinado siempre se solicita a un mismo servidor Web (origen) y sólo el sistema de encaminamiento de peticiones tiene la facultad de solicitarlo a otro servidor (réplica) si lo estima conveniente.

Las peticiones de los clientes se distribuyeron uniformemente entre el número de clientes modelados. Los contenidos también se distribuyeron uniformemente sobre los servidores Web del modelo y para poblar los distintos servidores réplicas se consideró replicar todos aquellos contenidos cuyo tamaño fuese mayor a 50KB y su demanda superior a 20 peticiones diarias. Estos contenidos fueron replicados en cada uno de los servidores réplicas del modelo y se generaron las tablas de ubicación respectivas. Nótese que estas decisiones tienen como único objetivo acelerar y facilitar la carga de datos del modelo de simulación. Claramente, si se conoce información más detallada de la red a modelar, se podría cambiar la distribución de peticiones en los clientes, la distribución de contenidos en los servidores Web y la política de doblamiento de los servidores réplicas.

Los resultados obtenidos de las simulaciones realizadas del esquema propuesto, se compararon con los resultados de simulaciones que se hicieron con un esquema de selección aleatoria y un esquema de selección Round Robin (RR) del servidor réplica. La comparativa queda reflejada en los gráficos que se presentan en la Fig. 6, cada medición corresponde a un valor promedio del tiempo de respuesta normalizado conseguido después de 5 simulaciones completas, cada una de las cuales abarca un periodo de una semana de requerimientos.

En cada nueva simulación se aprovecha la experiencia obtenida en las simulaciones anteriores, manteniendo los valores estimados de los tiempos de respuesta normalizados de la simulación previa.

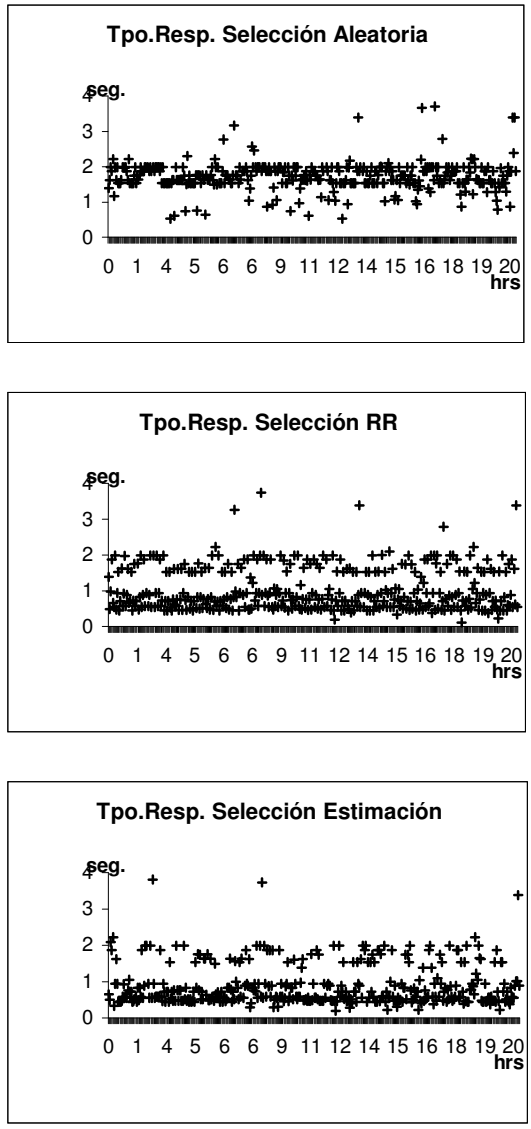


Figura 6: Tiempos de respuesta resultantes selección aleatoria, RR y por estimación.

El factor α se mantuvo constante en 0.5 durante todas las simulaciones, esta decisión responde al hecho del buen rendimiento obtenido en la primera simulación con respecto a la selección aleatoria, sin embargo, una cuestión pendiente es experimentar variando α.

Los resultados arrojaron que el esquema de selección de servidor réplica propuesto, mejora los tiempos de respuesta normalizados en cerca de un 20% en comparación a la utilización de un esquema de elección aleatorio de servidor réplica y en cerca de un 7% en comparación a la utilización de un esquema de selección RR. En el segundo caso, creemos que el escaso mejoramiento logrado en el rendimiento del esquema de selección, se puede deber a la localización equidistante de los servidores réplicas con respecto a la zona de usuarios y al hecho de que sólo se incluyó una zona de usuarios dentro del área en las simulaciones. Queda entonces por experimentar con distintas localizaciones de los servidores réplicas de un área específica e incluir más

de una zona de usuarios en ella. Además, aún nos falta por evaluar el rendimiento de este esquema de selección unido a un mecanismo de redirección.

5 Conclusiones y comentarios.

La aparición de las CDNs, sugiere vías de introducir nueva funcionalidad cerca de los extremos pero dentro de la red. Una arquitectura distribuida CDN podría implementarse como una red superpuesta en máquinas situadas en la periferia, entre el núcleo de la red y los usuarios finales (red overlay). Por otro lado, la implantación de estas soluciones no es trivial dada la complejidad de su arquitectura. Nosotros hemos planteado en este artículo una arquitectura adaptable para el sistema de encaminamiento de peticiones de una CDN, que puede perfectamente interactuar con cualquier sistema de distribución y sistema de tarificación de una CDN. Además de ser una propuesta que escala bien, refleja la percepción que tiene el usuario del rendimiento del sistema.

En concreto, nuestra propuesta de arquitectura permite realizar un esquema de medición del lado cliente en combinación con un criterio de selección estadístico. Esta combinación es novedosa y reúne lo mejor de ambos componentes. Por un lado, el esquema de medición del lado cliente permite reflejar de mejor forma la percepción que el usuario tiene del rendimiento del sistema. Eso unido a la métrica elegida, tiempo de respuesta final, sobre la cual el usuario tiene una gran sensibilidad, hacen que se tenga una visión más real de la experiencia que está viviendo el usuario con el sistema. Por otro lado, al aplicar un criterio de selección de tipo estadístico, se minimiza el overhead de tráfico que producen los criterios de selección dinámicos, ayudando a mejorar las condiciones del sistema.

Mediante simulación, utilizando NS-2, hemos comparado nuestra propuesta con una selección de servidor réplica aleatoria y RR, los resultados arrojaron que los tiempos de respuesta mejoran en casi un 20% y 7% respectivamente cuando se realiza la elección apoyada en la estimación estadística que proponemos.

Finalmente, hemos incorporado dentro de nuestra propuesta, parte de un mecanismo de adaptación que puede ayudar a aumentar las prestaciones de una CDN y que no provoca un mayor overhead en la red. Este mecanismo realiza las solicitudes de crear o eliminar copias de contenido según el comportamiento de la CDN, pero deja la decisión final al sistema de distribución con el cual interactúe.

Agradecimientos

Este trabajo ha sido financiado en parte por el Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica (I+D+I) y el Ministerio de Ciencia y Tecnología, a través del proyecto TIC2003-04406.

Referencias

[1] P.Rodríguez, C.Spanner and W.Biersack, "Analysis of Web Caching Architectures: Hierarchical and Distributed Caching", IEEE/ACM Transactions on Networking, vol.9, pp.404-418, August 2001.

[2] E.Turrini, "An Architecture for Content Distribution Internetworking". TR UBLCS-2004-2. March 2004.

[3] M.Day, B.Cain, G.Tomlinson and P.Rzewski, "A Model for Content Internetworking (CDI)", RFC 3466, Feb. 2003.

[4] G.Peng, "CDN: Content Distribution Network", TR-125, http://arxiv.org/abs/cs/0411069, Nov, 2004.

[5] H.Ossandón, E.Pastor. "Caché Web y Redes de Distribución de Contenidos: Una Visión General", JITEL2003, pp.551-552, España, Septiembre 2003.

[6] A.Babir, B.Cain, R.Nair and O.Spatscheck, RFC3568 "Know Content Network (CN) Request-Routing Mechanisms", http://www.rfc-archive.org/getrfc.php?rfc=3568, July 2003.

[7] S.Crespo y J.Rodríguez, "Redes de Distribución de Contenidos". Comunicaciones de Telefónica I+D, Nro 24, pp.193-204, Enero 2002.

[8] R.Mukhtar, Z.Rosberg. "A Client Side Measurement Scheme for Request Routing in Virtual Content Distribution Networks". Proc. IEEE International Performance, Computing, and Communications Conference, Phoenix, AZ USA , April 2003.

[9] Dykes, S. G., Robbins, K. A. and Clinton Jeffery, C. L., "An Empirical Evaluation of Client-side Server Selection Algorithms". Proc. INFOCOM'00, March 2000, pp. 1361-1370.

[10] M. Andrews, B. Shepherd, A. Srinivasan, P. Winkler and F. Zane, "Clustering and Server Selection using Passive Monitoring", Proc . INFOCOM'02, 2002.

[11] J.Liu, S.Yang, H.Yu and L.Tseng. "Content Delivery Network with Hot-Video Broadcasting and Peer-to-Peer Approach". In Journal of Information Science and Engineering 20, 1125-1139, Jun. 2004.

[12] P.Barford, J.Y.Cai, and J.Gast, "Cache placement methods based on client demand clustering" TR1437, University of Wisconsin / Madison, March 2002.

[13] A.Bestavros and S.Mehrotra. "DNS-based internet client clustering and characterization". Tech. Rep. BUCS-TR-2001-012, Boston University, 2001.

[14] P.Mehra, C.De Vleeschouwer and A.Zakhor, "Receiver-Driven Bandwidth Sharing for TCP," in *Proceedings of IEEE INFOCOM 2003*, 2003.

[15] "The Network Simulator NS-2: Manual", http://www.isi.edu/nsnam/ns/doc/ns_doc.pdf

[16] J.Chung and M.Claypool, "NS by Example", http://nile.wpi.edu/NS/, última visita Marzo/2005.

[17] E. Pinheiro, "Rutgers University Web Server logs", http://www.cs.rutgers.edu/~edpin/logs.html September, 2002.

Estudio y recomendaciones en el uso de protocolos de streaming en Redes Heterogéneas.[1]

Xavier Hesselbach i Serra, Joan Manuel Lopez Ruiz, Sergio Machado Sánchez
Departamento de Ingeniería Telemática. Universidad Politécnica de Cataluña
C/ Jordi Girona, 1 y 3 – Módulo C3 – Campus Norte.
08034 Barcelona
Teléfono: 93 401 59 87 Fax: 93 401 10 58
E-mail: {xavier.hesselbach, smachado}@entel.upc.edu

Abstract. *In this paper we study and discuss the behaviour of the streaming service offered in widely deployed commercial platforms, regarding the protocols and the information format. The study and trials are made in a real scenario, with congestion and traffic interruptions. This scenario includes access technologies such as WLAN and Bluetooth. These technologies are connected by means of an MPLS backbone network. This paper proposes some recommendations in order to enhance the compatibility and efficiency of the protocols and services running over heterogeneous networks among different platforms.*

1 Introducción

Las aplicaciones basadas en contenidos multimedia constituyen actualmente un gran porcentaje de la utilización total de las redes, debido al interés de los usuarios, a la mejora de capacidad en las redes de acceso, a la mayor potencia de los equipos cliente y a la aparición de gran número de aplicaciones multimedia y peer-to-peer (P2P).

En este artículo se analizan dos objetivos fundamentales. El primero es el estudio experimental del servicio de streaming a través de distintos tipos de red para elaborar un conjunto de recomendaciones de diseño y uso. Se consideran tecnologías de red de acceso inalámbricas y fijas, para determinar y cuantificar los comportamientos anómalos existentes, tales como pérdidas de paquetes, retardos, tiempos de traspaso o interrupciones de tráfico, entre otros. El segundo es la compatibilidad entre las distintas plataformas de streaming, de manera que un servidor alojar y distribuir archivos de otras plataformas.

Esta ponencia se organiza del siguiente modo: Se presenta en la sección 2 una revisión de conceptos fundamentales asociados al streaming de contenidos multimedia y de protocolos para su transporte. La sección 3 describe los escenarios de prueba, presenta una selección representativa del numeroso conjunto de medidas experimentales efectuadas y ofrece los resultados de los tests de compatibilidad entre las plataformas comerciales de Microsoft (Windows Media) y Real Media, en un escenario con interconexión entre redes diferentes, donde exista riesgo de congestión o interrupción de tráfico. A partir de estas medidas, en la sección 4 se recopilan las principales recomendaciones de uso práctico y diseño de futuros nuevos protocolos de streaming. La ponencia finaliza con las conclusiones finales.

2. Streaming

2.1 Streaming de Contenidos Multimedia

El concepto de streaming viene tradicionalmente asociado a contenidos multimedia. En realidad, streaming es un concepto más general, apto para el transporte de incluso información de señalización de red (como puede ser el caso del protocolo SCTP, Streaming Control Transport Protocol) [10], cuyo proposito de diseño fue el transporte de señalización de PSTN sobre redes IP, aunque ya se preveía la posibilidad de aplicaciones mucho más amplias.

De acuerdo al Open Source Streaming Alliance, streaming es el término empleado para describir la entrega en *tiempo real* de imágenes en movimiento, sonido e incluso texto, en una red de datos. La recuperación de archivos mediante streaming se caracteriza por proporcionar los datos al equipo cliente a medida que los necesita. Adicionalmente, el streaming debe ser capaz de proporcionar un camino de retorno hacia el servidor, por el cual permitir el envío de comandos desde el cliente para funciones de control del flujo. El término *tiempo real* debe interpretarse como la entrega de la información, con una acotación de retardo a un valor que permite al usuario tener noción de transmisión en directo.

La clave de este mecanismo está en el almacenamiento de los datos recibidos en un buffer, lo que permite contrarrestar el jitter en la transmisión de paquetes y reproducir el archivo a la tasa de

[1] Este trabajo ha sido parcialmente financiado por el proyecto nacional TIC2003-08129-C02.

codificación. El tamaño de este buffer se suele expresar en unidades de tiempo. Su valor por defecto en aplicaciones comerciales se establece entre 3 y 5 segundos.

2.2 Protocolos de Streaming

El establecimiento de una sesión para la recuperación de un archivo de vídeo se realiza tradicionalmente con alguno de los siguientes protocolos: Real-Time Streaming Protocol (RTSP) [1], HyperText Transfer Protocol (HTTP) [2], Real Time Protocol (RTP) [11] y Session Initiation Protocol (SIP) [3]. Otros protocolos, o bien son de uso menor, como SCTP, o bien han quedado obsoletos, como MMS (Multi Media Server, de Microsoft).

El protocolo RTSP implementa una máquina de estados finitos, lo que permite al usuario tener control sobre la reproducción del archivo, sin necesidad de cerrar la sesión establecida.

El protocolo SIP posee unas características que lo hacen adecuado para el establecimiento y el control de llamadas multimedia en entornos móviles.

La recuperación de datos mediante técnicas de streaming admite el transporte mediante los protocolos Transmisión Control Protocol (TCP) [4] y User Datagram Protocol (UDP) [5]. Al tratarse de una aplicación en tiempo real se requieren protocolos específicos de apoyo al transporte y control del flujo de datos, especialmente en el caso de UDP, ya que carece de ellos.

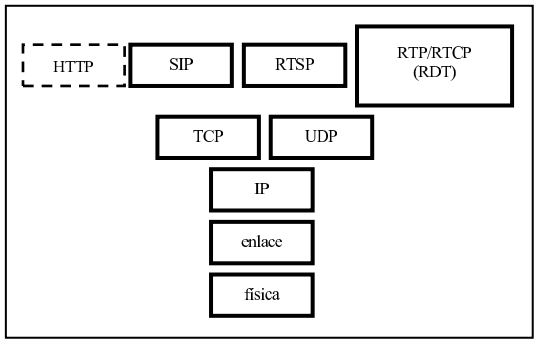


Figura 1: Torre de protocolos de streaming

Con este fin se implementan el protocolo Real-time Transport Protocol/Real-time Transport Control Protocol (RTP/RTCP) [6], estandarizado por el Internet Engineering Task Force (IETF) [7], y el protocolo RealNetworks Data Transport (RDT), propietario de Real Networks. En la Fig.1 se muestra la torre de protocolos relacionada a entornos comerciales de streaming.

Teniendo en cuenta la transmisión de contenidos de video mediante streaming, en general, son preferibles protocolos sin mecanismos de recuperación de errores (estilo UDP), puesto que el tiempo necesario para la corrección mediante intercambio de tramas convierte en caduca la información una vez ha sido corregida. Por ello, es preferible no corregir la información a menos que se efectúe por medio de técnicas FEC (Forward Error Correction), o de corrección en destino.

Por otro lado, hay que considerar el mecanismo de control de congestión existente en TCP, pero no implementado en UDP. En líneas generales, TCP se comporta reduciendo el caudal cuando detecta congestión en la red, pero no actúa de este modo UDP. En consecuencia, tal como se puede comprobar en pruebas experimentales, la compartición de caudal entre una conexión TCP y una UDP no es equitativa, quedando balanceada hacia UDP.

Por este motivo, teniendo en cuenta que TCP es un protocolo ampliamente usado en Internet, se recomienda el empleo de protocolos denominados *TCP-Friendly*, que incorporan mecanismos de control de congestión para que no hundan el caudal de las conexiones TCP.

3 Escenarios y pruebas

El método seguido en este estudio consta de tres partes: el análisis del comportamiento de las plataformas comerciales de streaming, el análisis de las situaciones anómalas previstas en las distintas tecnologías de red sobre las que se pretende ofrecer servicios de streaming y, por último, el análisis de estas tecnologías de red para caracterizar su comportamiento.

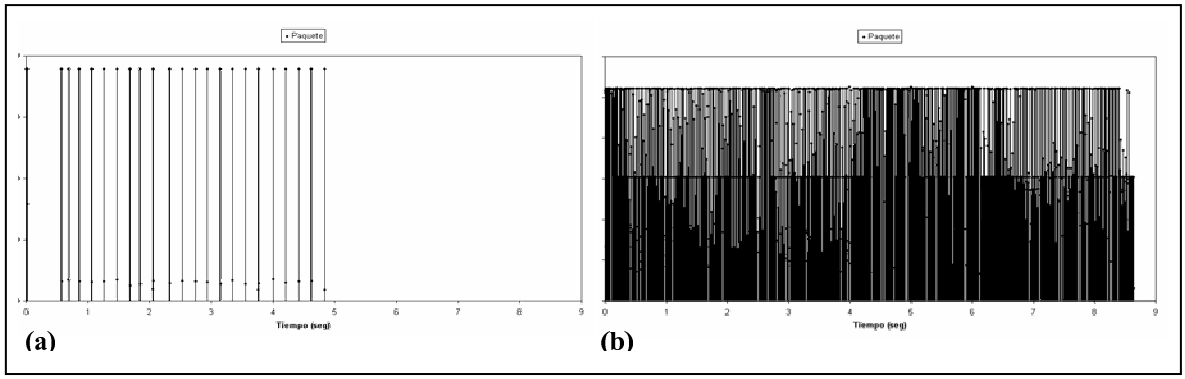


Figura 2: Envío de datos del servidor al cliente. **(a)** Windows Media. **(b)** Real Media

Las dos primeras partes se desarrollan sobre una red de área local Fast-Ethernet conmutada, compuesta únicamente por tres equipos, en la que se ubican el servidor de streaming y el cliente.

La tercera parte se desarrolla sobre cada una de las tecnologías de red bajo estudio: Bluetooth, Wireless LAN (WLAN) 802.11b y 802.11g, y Multi-Protocol Label Switching (MPLS).

Las pruebas experimentales se han realizado con las plataformas Windows Media 9 y Real Media 10 con servidor Helix y reproductor Real One, sobre sistema operativo Windows XP. En ambos casos se ha utilizado la misma fuente, un archivo de formato .avi, de 31 segundos de duración, codificado a 240 kbps CBR (Constant Bit Rate) y SBR (single Bit Rate) en el caso de Windows Media y a 200 kbps en Real Media, también de tasa constante única.

La codificación también podría haberse efectuado MBR, Multiple Bit Rate, en un único flujo, y también de tasa variable, pero no es necesario para los experimentos que se van a llevar a cabo.

3.1 Análisis de las Plataformas de Streaming

Se analiza en primer lugar el comportamiento de las dos plataformas de streaming, Windows Media [8] y Real Media [9]. Para ello, se han realizado repetidas pruebas, estableciendo sesiones mediante protocolo RTSP, con transporte TCP y UDP para la transmisión de un contenido de vídeo.

Ambas plataformas utilizan el protocolo TCP durante el proceso de establecimiento de la sesión con RTSP por cuestiones de fiabilidad. El uso que hacen de RTSP es distinto, aunque respetan el funcionamiento de la máquina de estados definida en [1].

Cuando transfieren el archivo al cliente, el servidor de Windows Media envía los datos a ráfagas, mientras que el servidor de Real Media los transmite a la cadencia correspondiente a la tasa de codificación del fichero, durante todo el tiempo que dure la transmisión. En la Fig. 2 se muestra este comportamiento.

La gráfica muestra cada trama como una línea vertical, cuya altura indica su tamaño en bytes. Se observa que el tamaño de los paquetes en la red es distinto en función del protocolo de transporte, en ambas plataformas.

Se ha observado que cuando el transporte se realiza por medio de protocolo TCP el servidor envía los datos en grupos. En Windows Media estos grupos son del orden de 30000 octetos mientras que en Real Media son de, aproximadamente, 2400 octetos. El resultado es que en red se observan ráfagas o agregados de paquetes Ethernet, cuyo valor responde a estos tamaños. Cuando el transporte es UDP, la fragmentación del archivo viene dada por las marcas que introduce el codificador en él. En Windows Media el tamaño del bloque es fijo, mayor que el de un paquete Ethernet, y dependiente de la velocidad de codificación. En Real Media los fragmentos son de tamaño variable y menores que el de un paquete Ethernet.

Si la selección en el reproductor del tipo de conexión disponible es adecuada (usualmente en el menú Herramientas-Opciones-Red del reproductor), se ponen en marcha los mecanismos de apoyo a streaming propios de cada plataforma, que básicamente, aceleran la velocidad del flujo de datos durante la descarga, dando lugar a gráficas de caudal instantáneo no esperadas a priori.

Windows Media configura mecanismos de aceleración de datos para el llenado del buffer inicial (Fast Start) y de aceleración del flujo de datos durante la descarga, éste último sólo si el transporte es TCP y se habilita el almacenamiento del clip en la carpeta de archivos temporales de Internet. Estos mecanismos se activan desde el Servidor de Windows Media.Real Media tan sólo dispone de un mecanismo de aceleración de flujo durante la descarga, que permite la transferencia a una velocidad hasta 4 veces mayor que la tasa de codificación del clip. En la Fig. 3 se muestra un ejemplo de este comportamiento para cada una de las plataformas. Estos mecanismos son útiles para garantizar el llenado del buffer del reproductor y garantizar la existencia de datos suficientes para evitar interrupciones de imagen.

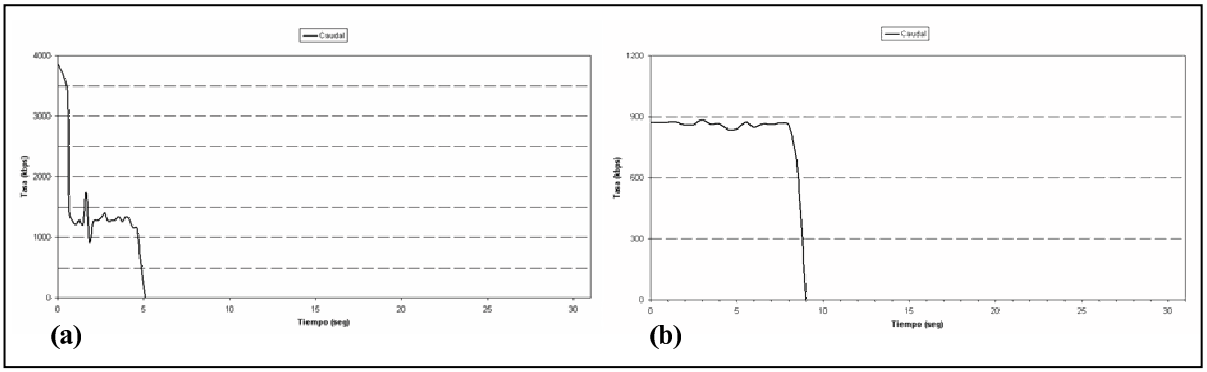


Figura 3: Descarga de un clip con los mecanismos de aceleración de flujo activados (a) Windows Media (b) Real Media

3.2 Soluciones de Compatibilidad Actuales

Las plataformas comerciales no son en general compatibles. Las causas se articulan en dos ejes: Falta de compatibilidad por protocolo y por formato de la información.

Se han estudiado los mecanismos que permitan que el servidor de un sistema sea capaz de alojar y distribuir archivos de otros. Actualmente, de las dos plataformas analizadas, sólo Real Media permite alojar y distribuir archivos de otras, en las versiones más avanzadas de sus servidores, como el Helix Universal Server.

En diversas pruebas y configuraciones se ha experimentado con la recuperación de un archivo de Windows Media alojado en el servidor de Real Media, pero los resultados obtenidos no han sido satisfactorios.

La descarga directa a través del reproductor se acaba negociando entre cliente y servidor mediante el protocolo HTTP, a pesar de tener habilitadas todas las opciones de RTSP y entrar la URL como rtsp:// y garantizar que no exista ningún cortafuegos limitando los puertos o servicios. La descarga a través de un enlace en una página web no llega a completarse, ya que el servidor pretende utilizar el protocolo Microsoft Media Server (MMS), propietario de Microsoft, y que los elementos de la última versión de Windows Media no implementan.

3.3 Streaming en Congestión

En esta sección se evaluará el comportamiento de los protocolos empleados para la transmisión de streaming en condiciones de congestión en la red. El estudio en entornos con congestión se ha efectuado limitando el caudal en la interfaz de red mediante el software Bandwidth Controller [12], con restricciones de caudal en periodos de tiempo determinados (de duración 5 segundos, con un caudal máximo permitido de 100 Kbit/s) de la transmisión de streaming, emulando tanto el caso de existencia de tráfico de fondo como de caudal reservado para otras aplicaciones.

El protocolo de transporte TCP no detecta de forma explícita la congestión, sino por vía indirecta considerando el número de pausas que se producen al tratar de conseguir el relleno del buffer. Aun así, gracias al mecanismo de reconocimiento de tramas recibidas, puede mantener el control de la transmisión. Con este mecanismo, en la reproducción del clip se observan pausas de imagen, pero la calidad es correcta para un usuario medio.

Por otro lado, el protocolo de transporte UDP no tiene ningún control sobre el flujo, por lo que necesita de los protocolos de apoyo al transporte y control de datos específicos de streaming, RTP/RTCP y RDT, para detectar la congestión. La reproducción del clip es continua, pero deficiente, como se pone de manifiesto en la representación de los movimientos.

La plataforma Windows Media dispone de mecanismos de actuación ante congestión, por los que el cliente solicita al servidor el cese de la transmisión del canal de vídeo pero no el de audio, con lo que se reproduce únicamente el canal de sonido hasta el final del clip. En la Fig. 4 se muestra un ejemplo, donde la congestión se presenta con la transmisión iniciada, finalizando tras unos segundos. El protocolo de transporte en este caso es UDP.

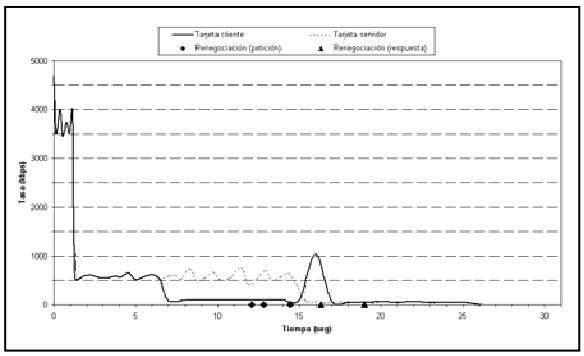


Figura 4: Actuación de Windows Media ante congestión

Como se ha constatado experimentalmente, Real Media no dispone de un mecanismo concreto de actuación ante congestión. En situaciones puntuales, el servidor puede reducir unilateralmente la tasa de transmisión, enviando sólo el canal de audio.

En el caso que desapareciera la congestión antes que la transmisión completa del clip, en el caso que se hubiera producido una renegociación de parámetros a la baja (por ejemplo, reducción del ancho de banda), éstos no se vuelven a negociar.

Así pues, se mantiene la descarga en las condiciones impuestas por la situación anómala, a pesar de que las nuevas condiciones son mejores. Esta situación debe ser solventada, y puede ser especialmente grave en redes inalámbricas, donde los procesos de traspaso pueden ser interpretados por los protocolos de transporte como situaciones de alta congestión en la red.

Una posible solución a los problemas observados durante la reproducción en situaciones de congestión, es el uso de clips codificados a Multiple Bit-Rate (MBR). Se observa su eficacia en la negociación de parámetros en el proceso de establecimiento, gracias a los mecanismos propios implementados por cada plataforma, pero en las pruebas realizadas en el laboratorio se ha observado, en contra de lo esperado a priori, es una solución no siempre válida si la congestión se inicia con la descarga ya iniciada, debido al retardo en el acuerdo de la nueva tasa que muchos usuarios no estarían dispuestos a aceptar.

3.4 Streaming ante Interrupciones de Tráfico

La detección correcta de interrupción requiere un mecanismo de realimentación entre el servidor y el cliente. Este mecanismo puede ser por reconocimiento de tramas de TCP. La plataforma Real Media implementa su propio mecanismo junto al protocolo RDT.

Windows Media dispone de soluciones de retransmisión tras el corte, que permiten seguir visualizando el clip desde el punto en que se interrumpió la comunicación. El servidor de Windows Media es capaz de esperar el reestablecimiento de la comunicación con intervalos de tiempo indefinidos, a pesar de agotar los tiempos de espera del mecanismo de back-off exponencial de TCP.

Si el protocolo de transporte es UDP, el corte no se detecta hasta el final de la transferencia del archivo, cuando el servidor envía un mensaje de RTSP en que anuncia este final. Como este mensaje va sobre TCP, requiere un reconocimiento por parte del cliente, que, en caso de mantenerse la interrupción, el servidor no recibe. Si esta situación finaliza antes de completarse la descarga del clip, el cliente espera a que el servidor anuncie el final para solicitarle la retransmisión.

Esta situación se muestra en la traza de caudal mostrada en la Fig. 5. Se observa que al inicio de la retransmisión se envía una ráfaga de tramas de forma muy próxima en el tiempo, (las líneas están más juntas). Esto se debe a una técnica propia de Windows Media denominada *Fast Reconnect*, que acelera el flujo hasta llegar al punto en que se debe retomar la reproducción, facilitando el rápido llenado del buffer del reproductor.

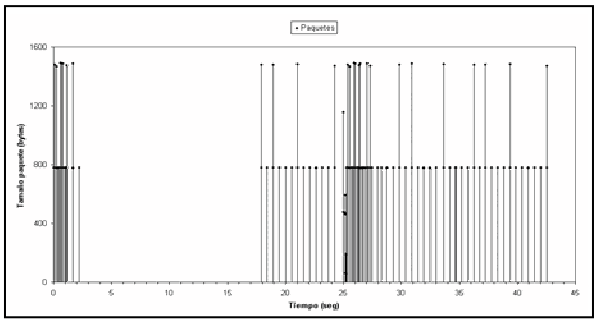


Figura 5: Interrupción del tráfico durante la descarga de un clip. Plataforma Windows Media, transporte UDP

Real Media no adopta ningún mecanismo de actuación ante la interrupción. Si el servidor detecta la anomalía, a partir de la no recepción de paquetes de realimentación desde del cliente, se cierran los canales de comunicación de forma inmediata.

3.5 Streaming sobre Bluetooth

Bluetooth es una tecnología inalámbrica de corto alcance, pensada para interconexión cercana de equipos a tasas moderadas (centenares de Kbit/s). Bluetooth se caracteriza por tener un caudal decreciente al aumentar la distancia entre dispositivos (Fig. 6). Este caudal debe ser compartido por todos los dispositivos conectados. Es una tecnología válida para la transferencia de vídeo streaming, aunque de forma muy limitada debido a sus características, tanto por capacidad como por alcance de los destinatarios, aunque puede ser muy útil en museos y otros espacios similares.

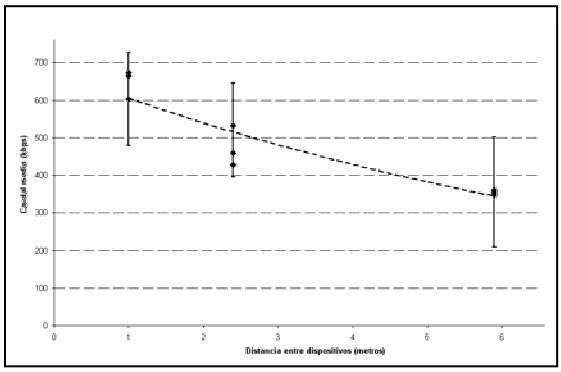


Figura 6: Tecnología Bluetooth. Evolución del caudal medio en función de la distancia

En las conexiones de streaming mediante Bluetooth son habituales las situaciones de congestión, ya que el ancho de banda disponible para este tipo de aplicaciones es más bien escaso, unos 550 Kbps a 2 metros, especialmente si se conecta más de un dispositivo al servidor, ya que deben compartir este caudal.

En el laboratorio se han hecho diversas pruebas de streaming sobre esta tecnología, que incluyen la transmisión de un clip por parte de un único cliente, y la descarga de varios clips, secuencial y simultáneamente, por parte de hasta tres clientes.

En las pruebas con un cliente, cuando la tasa de codificación es la mitad de la disponible teórica, en función del resultado obtenido en la Fig. 6, la reproducción es correcta. Si se alejan los dispositivos, provocando la disminución del caudal disponible, la reproducción se asemeja a la observada en los casos en que se produce congestión.

En las pruebas con tres clientes, se comprueba que se reparten el caudal disponible, por lo que si la suma de las tasas de codificación de los clips solicitados, es cercana por defecto o superior a la disponible, aparecen los efectos debidos a la congestión. En la descarga secuencial de clips, se observa que a partir de solicitar el segundo, el primero empieza a presentar anomalías en la reproducción, que se hacen más evidentes al solicitar el tercero. En el laboratorio se ha comprobado que si la suma de los clips

solicitados es menor al 60% del caudal disponible, no se observan anomalías en la reproducción.

3.6 Streaming sobre WLAN

Actualmente se está dotando de infraestructuras WLAN lugares tales como aeropuertos, estaciones de ferrocarril, hoteles o cafeterías, que facilitan a sus usuarios y clientes el acceso a Internet. El aspecto más crítico a la hora de ofrecer servicios de vídeo streaming a través de WLAN, es el tiempo de roaming, es decir, el tiempo que necesita un terminal en conectarse a un Access Point cuando sale de la zona de cobertura de otro que le daba servicio hasta ese momento. Durante el roaming, se produce la interrupción del tráfico, ya que se trata de un traspaso del tipo hard-handover. En el laboratorio se ha medido este tiempo, y el resultado ha sido 8,5 segundos.

WLAN 802.11 puede funcionar en modo Ad-Hoc o infraestructura (basado en un Punto de Acceso). Para las pruebas que se presentan se han realizado experimentos en el segundo caso, empleando el estándar 802.11b, con un caudal teórico de 11 Mbps, pero un caudal neto de poco más de 9 Mbps.

El ancho de banda disponible en esta tecnología de acceso, aun siendo compartido por todos los terminales que acceden a un Access Point, es suficiente para ofrecer servicios de streaming de contenidos multimedia. En la Fig. 7 se muestra la densidad de probabilidad de caudal en la comunicación entre dos tarjetas de red, en modo infraestructura, siguiendo las especificaciones del estándar 802.11b. La tasa media obtenida, del orden de 4700 kbps, demuestra que ambos dispositivos comparten el caudal disponible (hasta 11 Mbps).

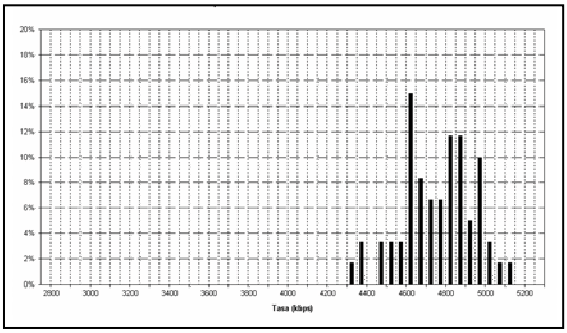


Figura 7: Función densidad de probabilidad de caudal en la comunicación de dos tarjetas WLAN en modo infraestructura

3.7 Streaming a través de MPLS

La tecnología de red MPLS se caracteriza por la capacidad de ofrecer políticas de calidad de servicio, por servicios diferenciados, al tráfico que la atraviesa. Una de estas políticas hace referencia al ancho de banda máximo que se otorga a un cierto tipo de tráfico, y que sólo puede sobrepasar un cierto exceso definido por el operador.

En el laboratorio se ha transmitido el clip de la plataforma Windows Media, codificado a 240 kbps, con un pico de caudal de salida inicial debido al mecanismo de Fast-Cache de 1500 kbps, a través de un Label Switched Path (LSP) limitado a 400 Kbps mediante el mecanismo de MQC (Modular QoS CLI) [13] que incorporan los routers Cisco.

En los diversos experimentos efectuados, se han observado anomalías al inicio de la reproducción, similares a las producidas ante congestión. Estas anomalías se recuperan cuando el servidor finaliza el pico de tráfico inicial, e continúa la transmisión a una tasa de unos 200 Kbps, inferior al límite impuesto, como se observa en la Fig.8.

Desde el punto de vista del servicio de vídeo streaming, como los observados en este estudio, esta característica puede ser problemática en el caso de los flujos acelerados de salida, ya que se supera ampliamente, por unos momentos, el ancho de banda asignado a este servicio. Hay que tener muy en cuenta este aspecto, porque se supone que los operadores de redes MPLS intentarán restringir al máximo el uso de recursos de la red que hagan sus clientes.

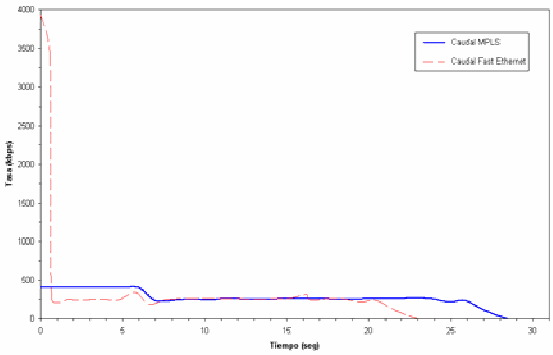


Figura 8. Streaming al atravesar un backbone basado en MPLS con gestión de caudal mediante MQC.

4 Recomendaciones

Esta sección recoge los principales resultados extraídos del conjunto de experimentos realizados empleando las diversas tecnologías de red, de forma independiente como interconectadas entre sí.

En relación al establecimiento de sesiones, debe tenerse presente que:

Como consideración general, la compatibilidad entre plataformas se puede garantizar usando los protocolos estándar definidos por el IETF en los procesos de establecimiento y transferencia.

Una vez el cliente toma contacto con el servidor, la negociación de los parámetros básicos para la transferencia debe hacerse en las peticiones y respuestas de los métodos RTSP que afectan a la máquina de estados. En el caso de la transferencia

mediante UDP, se deben dar a conocer las marcas que delimitan los paquetes, para facilitar la fragmentación.

En relación a los formatos y su compatibilidad, los entornos actuales, por ser propietarios, no facilitan la distribución de contenidos de su propia plataforma. Esto es más evidente en Windows Media que en Real Media. En este último, existe un cierto grado de admisibilidad de formatos, que son servidos si se llega a un acuerdo de protocolo de transporte con el cliente.

En relación a la gestión de recursos y a los protocolos:

Ante situaciones de congestión se debe aprovechar la capacidad renegociadora de RTSP, y la posibilidad de crear archivos MBR, para habilitar la descarga del clip a una tasa menor. Para facilitar el control de la tasa de descarga en estos casos, y evitar la pérdida de tramas, se debe dotar al protocolo de control de flujo RTCP de la capacidad de reconocer al servidor las tramas que recibe, de forma similar a como lo hace TCP, aunque sin permitir retransmisiones.

Es necesario crear un mecanismo de realimentación entre el cliente y el servidor que les permita conocer en todo momento que la conexión sigue establecida, especialmente en el caso del transporte UDP, que carece de este tipo de mecanismo. Asimismo, es necesario establecer un tiempo de espera máximo para la reconexión, suficientemente grande para afrontar situaciones de roaming en WLAN, pero facilitando que los equipos puedan cerrar los canales de comunicación de forma unilateral si se sobrepasa.

En consecuencia, debe modificarse el protocolo RTP/RTCP de manera que envíe al servidor mensajes que habiliten la salida de nuevos datos, a la vez que cumplen con la función de feedback, al estilo de TCP.

A partir del análisis de las tramas intercambiadas durante el establecimiento de sesión empleando protocolo RSTP, se sugiere mejorar de este protocolo aplicando los siguientes criterios: Deben negociarse los parámetros básicos para la conexión en los métodos que afectan la máquina de estados de RTSP, ya que siempre aparecen y tienen un uso especificado. El formato de las cabeceras y variables asociadas a estos parámetros básicos debe ser el mismo en todas las plataformas, para facilitar el entendimiento entre elementos de distintas plataformas

Para evitar problemas en la reproducción, debidos a un excesivo uso de ancho de banda, se debe seleccionar de forma manual el tipo de conexión disponible, de manera que sea lo más próximo posible, por exceso, a la tasa de codificación del clip.

Se deben usar archivos codificados MBR para favorecer la correcta transferencia mediante streaming ante situaciones de congestión y facilitar la adaptación al caudal disponible en la red en todo momento para maximizar la calidad percibida por los usuarios. En esta situación, en el caso de producirse deficiencias en la reproducción (usualmente debido a estar transmitiendo un contenido de tasa mayor que el ancho de banda disponible), el servidor debe optar por descargar una codificación de menor tasa del mismo archivo MBR.

Ante situaciones de interrupción, se debe implementar un mecanismo de feedback que facilite la detección, especialmente en el caso del transporte UDP. Si se produce una interrupción en la conexión no tiene sentido que el servidor siga enviando datos al cliente. Se debe implementar un tiempo de espera que permita mantener la conexión abierta en caso de pequeñas interrupciones como las provocadas por los traspasos en tecnologías de acceso inalámbricas

5 Conclusiones

En los análisis experimentales desarrollados en este estudio se ha comprobado el comportamiento de la transferencia de contenidos multimedia mediante streaming, cuando se establece la sesión con el protocolo RTSP, en las plataformas comerciales Windows Media y Real Media, sobre sistema operativo Windows. La transmisión se ha realizado empleando diversas tecnologías de red, interconectadas entre sí, con red troncal MPLS y de acceso de tipos Ethernet, WLAN y Bluetooth. De ello se ha derivado un prototipo experimental de pruebas para tráfico de streaming de propósito general.

Si la opción de conexión elegida es suficientemente general, ambas plataformas ponen en marcha mecanismos que aceleran el flujo durante la descarga.

En general, son preferibles soluciones basadas en UDP por motivos de retardo y ligereza del protocolo, aun cuando ello implica que debería añadirse soluciones TCP-friendly para evitar los inconvenientes de la convivencia de TCP con UDP desde el punto de vista de la repartición de caudal en los enlaces.

Las soluciones de compatibilidad entre elementos de distintas plataformas en la actualidad, son ineficientes. En el único caso en que se ha podido alojar un archivo de una plataforma en el servidor de la otra, la descarga sólo es parcialmente compatible. Aun así se observa cierta similitud en la forma de hacer el establecimiento con RTSP en ambas plataformas, lo que se puede aprovechar para crear una interfaz de adaptación entre ellas. Este resultado pone de manifiesto la tendencia de los diversos fabricantes a desarrollar soluciones de streaming particulares en buena medida.

La irrupción de nuevas tecnologías de acceso inalámbricas, como Bluetooth y WLAN, así como la progresiva implantación de redes de transporte con tecnología MPLS, hacen necesario considerar qué ocurre con la transmisión de un archivo de audio y/o vídeo mediante streaming, ante situaciones como congestión e interrupción de tráfico. Los resultados obtenidos en las pruebas realizadas ante este tipo de situaciones, demuestran que es necesario mejorar el comportamiento de los protocolos de transferencia para reducir la sobrecarga por señalización así como mejorar los procedimientos de adaptación de caudal al medio, durante la transmisión del contenido multimedia.

Los trabajos derivados de este estudio se centran en el desarrollo de una interfaz adaptadora entre protocolos y formatos (estilo *proxy*) de plataformas diferentes para su correcta comunicación. Por otro lado, en el estudio de los ajustes necesarios en el protocolo RTSP para conseguir dicha compatibilidad especialmente en el proceso de establecimiento. En la misma línea de los protocolos, las experiencias llevadas a cabo ofrecen un conjunto de condiciones que debería cumplir un protocolo robusto de streaming. Con ellas, los trabajos conducen a establecer un diseño TCP-friendly compatible con el resto de protocolos, para ser probado en condiciones reales.

Referencias

[1] H. Schulzrinne, A. Rao, R. Lanphier. "Real Time Streaming Protocol (RTSP)". IETF. Request for Comments (RFC) 2326. Abril 1998.

[2] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, T. Berners-Lee. "Hypertext Transfer Protocol HTTP/1.1". IETF. RFC 2068. Enero 1997.

[3] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler. "SIP: Session Initiation Protocol". IETF. RFC 3261. Junio 2002.

[4] Information Sciences Institute, University of Southern California. "Transmission Control Protocol". IETF. RFC 793. Septiembre 1981.

[5] J. Postel. "User Datagram Protocol". IETF. RFC 768. Agosto 1980.

[6] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson. "RTP: A Transport Protocol for Real-Time Applications". IETF. RFC 3550. Julio 2003.

[7] The Internet Engineering Task Force: http://www.ietf.org

[8] Microsoft Windows Media Series 9: http://www.microsoft.com/windows/windowsmedia/default.aspx

[9] Real Networks. Products: http://www.realnetworks.com/products/

[10] RFC 2960 "Stream Control Transmission Protocol". http://rfc.net/rfc2960.html. Octubre 2000.

[11] RFC 1889 "A Transport Protocol for Real-Time Applications". http://rfc.net/rfc1889.html. Enero 1996.

[12] Software de Control: Bandwidth Controller. http://www.bandwidthcontroller.com.

[13] Cisco Systems. MQC: "Modular Quality of Service Command-Line Interface". http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt8/qcfmdcli.htm

Ubicación eficiente de servicios de detección de pérdidas para VoIP en redes programables y de recubrimiento

Juan J. Ramos Muñoz y Juan M. López Soler
Departamento de Teoría de la Señal, Telemática y Comunicaciones.
Universidad de Granada.
C/ Periodista Saucedo Aranda s/n. 18017 - Granada
Teléfono: 958 24 23 03 Fax: 958 24 08 31
E-mail: jjramos@ugr.es, juanma@ugr.es

Abstract *In this work we present an algorithm for the efficient allocation of the packet loss detection service in an VoIP environment. It is well known that the final perceived quality will depend on the optimal design of all the services involved along the path. The main contribution in this work is the inclusion of a criterion based on a subjetive measure of the voice quality (by means of the use of a simplified version of the E-model) in the algorithm which estimates the detection agent location within a programmable or overlay network. To achieve this, the analytical expressions for the prediction of the loss probability and the mean delay are developed. Then, the proposed procedure is evaluated and validated by means of simulation.*

1. Introducción

Las aplicaciones de VoIP y, en general, cualquier aplicación que requiera la entrega de datos dentro de un intervalo de tiempo acotado, tienen grandes dificultades en su implementación sobre el servicio *best-effort* que IP proporciona. Entre otras cuestiones, el servicio no fiable proporcionado por IP no garantiza una entrega ordenada de paquetes, ni siquiera libre de errores, con control de flujo o libre de congestión. A resultas, las aplicaciones de VoIP, aun siendo en cierta medida tolerantes a fallos, demandan una Calidad de Servicio que no siempre las redes subyacentes interconectadas con IP pueden proporcionar.

Entre otras cuestiones, la recuperación de pérdidas de paquetes, aun no siendo un factor crítico, puede afectar de forma determinante en la calidad de voz extremo a extremo ofrecida. Las pérdidas son especialmente determinantes en escenarios donde las mismas se generen a ráfagas, es decir, entre paquetes consecutivos. Considerando las demandas de retardo extremo a extremo acotado que VoIP exhibe, es crucial, por tanto, minimizar los tiempos involucrados en la recuperación de las mismas. La recuperación de pérdidas puede llevarse a cabo extemo a extremo, pero por razones obvias de retardo, no es con mucho la mejor aproximación. Como alternativa, para llevar a cabo una detección temprana de pérdidas, es interesante la consideración de los servicios añadidos que las tecnologías de redes activas, programables o de recubrimiento (*overlay networks*) pueden ofrecer. En este caso, mediante el uso de memorias caché puede llevarse a cabo la recuperación en puntos intermedios reduciendo así el retardo extremo a extemo involucrado.

Aunque el uso de redes programables no requiere que todos los routers sean colaborativos (es decir, que sean programables) sino que por el contrario es suficiente su implantación en un subconjunto de ellos, esta tecnología puede encontrar grandes dificultades en su despliegue, tanto por objeciones relativas a la seguridad como por la posible merma y dificultad computacional que puede suponer para el router no sólo enrutar sino además procesar los datagramas. Alternativamente, la tecnología de redes punto-a-punto (P2P, *peer-to-peer*) está teniendo una amplia acogida entre usuarios finales, permitiendo la creación de extensas redes de intercambio de ficheros y, recientemente, la aparición de redes dedicadas a la transmisión de voz [1], utilizando como procesadores intermedios a puntos finales.

En VoIP, para paliar los efectos no deseados de las conocidas ráfagas o interrupciones en el flujo de información, antes de la recuperación misma de las pérdidas es necesario llevar a cabo su detección de forma eficiente (es decir, sin incurrir en excesivos retrasos). En [2] se proporciona un estudio detallado de posibles soluciones a este problema proponiendo y evaluando detectores de pérdidas adaptables a la dinámica de la red.

Para completar el diseño eficiente del sistema de recuperación de pérdidas de paquetes en VoIP, independientemente del detector de pérdidas adoptado e incluso independientemente del procedimiento o protocolo de recuperación in-

volucrado, es de especial relevancia, por su influencia en la calidad final resultante, la localización del servicio. La localización eficiente de servicios no es un tema nuevo, y ha sido abordado con rigor y éxito en trabajos previos para el contexto de redes activas. Por ejemplo, en [3] se propone un procedimiento basado en la activación y desactivación del servicio de recuperación basándose en la superación de un umbral superior e inferior de pérdidas respectivamente. No obstante, éste y otros trabajos relacionados exhiben claras limitaciones para su consideración en aplicaciones con demandas de tiempo real. Además, en la decisión de la localización no se tienen en cuenta criterios de calidad subjetiva, sino que tan sólo se tienen en cuenta parámetros que caracterizan el estado de la red.

En este trabajo se diseña un algoritmo que selecciona de forma eficiente la localización del detector y recuperador de pérdidas intermedio en base a un criterio de calidad subjetivo. Para ello, como se explicará más adelante, se adopta una simplificación del E-model. Bajo esta simplificación, estimando la probabilidad de pérdidas así como el retardo medio previsto, se puede obtener una medida analítica de la calidad de voz subjetiva resultante en términos MOS (*Mean Opinion Score*). Por tanto, dado un escenario, y un conjunto de posibles localizaciones, se proporciona un algoritmo sencillo que a partir del modelado analítico propuesto, permite la estimación de la calidad que se obtendría potencialmente para cada posible localización, y en definitiva la localización eficiente del servicio.

Para todo ello el trabajo se ha organizado en los siguientes apartados: en el apartado 2 se describe el modelo utilizado para medir la calidad de la solución aportada. A continuación, en la sección 3 se explica brevemente los mecanismos de recuperación con los que trabaja el algoritmo de localización. Más tarde se discute la estimación del valor MOS en un escenario dado de forma analítica, en el apartado 4. En las secciones 5 y 6 se describe el algoritmo y se evalúa su eficacia respectivamente. Por último las conclusiones se presentan en el apartado 7, y el trabajo a realizar en el 8. Por último se incluye una breve reseña bibliográfica de la literatura utilizada.

2. Medida de calidad de voz: E-model

En el desarrollo de servicios VoIP, evidentemente, un objetivo de diseño es maximizar la calidad de la voz transmitida. Para ello, es necesario disponer de medidas objetivas de evaluación. Es éste un problema complejo al que se le han dedicado muchos recursos, si bien dada su gran dificultad (intervienen tanto factores físicos como psicológicos), todavía no está definitivamente resuelto. No obstante, organismos internacionales de estandarización (ITU, TIA y ETSI) ya han especificado normas, por lo general, bastante aceptadas tanto para la estimación subjetiva (tests de opinión) como objetiva (modelos analíticos).

Aunque inicialmente concebido como una herramienta para la planificación de servicios, el así denominado E-model, especificado en la recomendación G.107 de la ITU-T, debido a su posible implementación con bajo coste computacional, está siendo utilizado para la medida o estimación objetiva de la calidad de la voz transmitida. Dado un entorno, E-model proporciona una estimación numérica ($R \in [0, 100]$) de la calidad a partir de la consideración de un elevado número de posibles dificultades en la transmisión (del orden de 20, por ejemplo, eco, ruido de fondo, pérdidas, degradación del codificador, retardos, etc). Todas ellas son modeladas analíticamente mediante 4 factores (I_s, I_d, I_e, A). El denominado *simultaneous factor* (I_s) está relacionado con la SNR del canal de transmisión. I_d es el *delay factor* que, como su nombre indica, modela todos los aspectos del retardo. I_e (el *equipment factor*) incluye todas las degradaciones introducidas por los codificadores y pérdidas de paquetes. Finalmente, en la estimación se incluye A (el *expectation factor*) como una corrección que modela la tolerancia que los usuarios están dispuetos a asumir por utilizar las ventajas (por ejemplo la movilidad) que proporcione el servicio. En definitiva, dado un entorno de transmisión de voz, se define

$$R = 100 - I_s - I_d - I_e + A \qquad (1)$$

A partir de R es posible estimar analíticamente, mediante la correspondiente conversión, el índice MOS (Mean Opinion Score)(véase la recomendación P.800 de ITU-T). MOS es una estimación subjetiva ampliamente aceptada consistente en un test de opinión o encuesta formulada a un gran número de oyentes (en general especializados) sobre un *corpus* de voz completo. Usando la estimación MOS a partir de E-model se evitan los gandes costes que supondría la realización de un MOS original, a la vez que permite la inclusión *on-line* de criterios de calidad en el desarrollo de los servicios de transmisión.

Dado que la aplicación de interés en este trabajo es VoIP, para la estimación de R se adoptarán las simplificaciones justificadas en [4]. Por tanto, si consideramos la función escalón $H(x)$, definida como

$$H(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{para } x >= 0 \end{cases} \qquad (2)$$

El factor R se puede expresar como

$$R \approx 94{,}2 - 0{,}024d + 0{,}11(d-177{,}3)H(d-177{,}3) - I_e \quad (3)$$

Donde d es el retardo extremo a extremo expresado en milisegundos. El cálculo de I_e depende entre otros factores del codificador utilizado, así como del patrón o distribución de las pérdidas de paquetes. Suponiendo un codificador PCM G.711 con ocultamiento de errores y un patrón de pérdidas aleatorio, la expresión para I_e puede estimarse como

$$I_e(G.711, random) \approx 30ln(1 + 15e) \quad (4)$$

En la expresión (4), e se define como la probabilidad de pérdidas total. Por último, la conversión del factor R (expresión 3) en el correspondiente valor sujetivo de calidad expresado en escala MOS se lleva a cabo mediante la siguiente expresión

$$MOS = 1 + 0{,}035 \cdot R + R \cdot (R-60) \cdot (100-R) \cdot 7 \cdot 10^{-6} \quad (5)$$

Es interesante resaltar que E-model proporciona una estimación de la calidad con una precisión limitada, ya que como se ha mencionado, por un lado el problema es de una gran complejidad, dado que está afectado por factores psicológicos siempre difíciles de modelar. Además E-model fue diseñado fundamentalmente para resolver problemas de planificación y no de evaluación propiamente dichos. De hecho, E-model está actualmente en revisión por el *Grupo de Estudio* 12 de la ITU-T. No obstante, por su sencillez, aun asumiendo cierta imprecisión, la inclusión de E-model en cualquier algoritmo para la provisión de un servicio implica un avance significativo, ya que de esta forma estaremos considerando (e incluso maximizando) un criterio de calidad de voz, objetivo final en la provisión del servicio.

3. Mecanismos de recuperación en redes activas

La introducción del uso de las redes activas en el ámbito de la recuperación de pérdida de paquetes ha dado como resultado distintas propuestas muy prometedoras [5]. Para reducir el retardo de la reparación y ofrecer soluciones escalables, se llevan a cabo acciones dentro de la red que agilicen el proceso de recuperación.

El mecanismo más ampliamente utilizado es el de la retransmisión local [3],[6]. Especialmente efectivo en árboles de multidifusión, este mecanismo consiste en implementar una memoria temporal en un nodo intermedio, que retransmite los paquetes solicitados por los receptores que hayan detectado alguna pérdida.

Para reducir aún más el retardo requerido en la recuperación de un paquete perdido se introduce otro mecanismo adicional, integrado con los anteriores, que consiste en realizar la detección en el nodo intermedio, efectuando la solicitud del paquete perdido al retransmisor más cercano.

Pero en las propuestas realizadas al respecto ([3], [7], [8]) la detección se basa en la observación de los números de secuencia de los paquetes. Se notifica una pérdida (y se reacciona con la retransmisión correspondiente) si se recibe un paquete cuyo número de secuencia no sigue al inmediatamente anterior, suponiendo de este modo que han sido perdidos los paquetes intermedios.

Sin embargo, este tipo de detección, que denominaremos DAS de ahora en adelante, adolece de una dependencia directa en la reacción en el caso de ráfagas de pérdidas. Cuanto mayor sea la longitud de una ráfaga, más tarde se producirá la detección.

Para que este servicio ofrezca una calidad adecuada minimizando los recursos de la red que requiere, es necesario identificar una ubicación para el servicio, de forma que se optimice la calidad percibida por los receptores del flujo de audio.

Este asunto ha sido abordado desde el punto de vista de la minimización de la probabilidad de pérdidas experimentadas extremo a extremo tras la recuperación [3]. No obstante, éste no es un parámetro suficientemente representativo para medir la calidad del flujo final en los receptores.

Para paliar estas deficiencias se proponen aproximaciones basadas en la evaluación del retardo de los paquetes, prediciendo el momento en el que deberían llegar, e identificando como pérdida los eventos en los que el paquete no llega en el tiempo estimado. En [9] se proponen varios ejemplos de detectores basados en retardos. En el apartado 3.1 se presenta un nuevo detector basado en retardos que utiliza el predictor de tendencias de Holt [10] para estimar la demora esperada para cada paquete.

El mecanismo de recuperación que se utilizará en este estudio se compone de los servicios de almacenamiento temporal, retransmisión local y detección temprana de pérdidas. En nuestro estudio se contemplarán dos módulos distintos de detección. El escenario a estudiar se ilustra en la Fig. 1, y el mecanismo descrito se denominará NRA. El nodo emisor S envía al receptor R paquetes que pasan por un nodo intermedio programable NRA, donde se ejecutan los mecanismos de recuperación. Entre los nodos S y NRA hay una probabilidad de pérdida p_1, y un retardo

de propagación igual a tp_1. Entre NRA y R hay una probabilidad de pérdida p_2, y un retardo tp_2.

Figura 1: Esquema del escenario básico estudiado.

3.1. Detección de pérdidas basada en predictor de retardos mediante Holt

En esta sección se propone un nuevo módulo de detección basado en un predictor de retardos para los paquetes pertenecientes a un flujo multimedia. Dicho módulo lo denominaremos DETAH (*Detección por Expiración de Tiempo Adaptable-Holt*). Como se ha descrito antes, su funcionamiento es sencillo: para cada paquete se estima el momento de llegada al nodo activo, de forma que si éste no ha sido recibido en ese plazo, se identifica como paquete perdido. Dada la existencia de fluctuaciones en el retardo de los datagramas en la red (*jitter*), es necesario que esta estimación se adapte al comportamiento de la red.

La justificación del diseño de este detector se centra en la gran correlación existente entre los retardos de paquetes adyacentes. De hecho podemos expresar dicho retardo como la combinación de un componente de tendencia y otro de ruido. Con esta suposición, proponemos un procedimiento de suavizado exponencial basado en el predictor de Holt. Concretamente, para un paquete i, la estimación del plazo en el que se espera su llegada $U_a(i)$ se calcula como:

$$\begin{aligned} L(i) &= \alpha d(i-1) + (1-\alpha)(L(i-1) + s(i-1)) \\ s(i) &= \beta(L(i) - L(i-1)) + (1-\beta)s(i-1) \\ D(i) &= L(i) + s(i) \end{aligned} \quad (6)$$

Donde $L(i)$ denota el nivel del dato, y $s(i)$ indica la estimación de la pendiente. Ya que puede siempre esperarse una componente de ruido, la predicción de la tendencia se modifica mediante el suavizado de $(L(i)-L(i-1))$ (la tendencia del último paquete) con β, y añadiéndosela a la estimación del retardo anterior $s(i-1)$, ponderada por $(1-\beta)$.
Para obtener una cota superior de la señal de retardo, basado en el suavizado exponencial de la ecuación (6), el umbral adaptable $U_a(i)$ se calculará finalmente como:

$$U_a(i) = D(i) + C \quad (7)$$

Donde la variable C se añade para reducir el número de falsas alarmas (errores en la predicción prematura) debido a la fluctuación aleatoria del retardo. Una vez que el paquete i ha llegado al nodo del detector, C se actualiza de la siguiente manera:

$$\begin{aligned} e(i) &= D(i) - d(i) \\ \delta(i) &= (1-\gamma)\delta(i-1) + \gamma\,|e(i)| \\ C &= \delta(i)\phi \end{aligned} \quad (8)$$

Las constantes que aparecen en las expresiones (6) y (8) (α, β, γ y ϕ) se definen en el rango $[0,1]$. Se utilizan para ponderar la influencia de los valores antiguos sobre los recientes.

Como nota final, en el caso de que el paquete previo se pierda, las ecuaciones (6), (7) y (8) no pueden evaluarse, y se utilizarán como alternativa los últimos valores estimados de C y D. En concreto:

$$U_a(i) = D(i-1) + C \quad (9)$$

De esta forma se obtiene un predictor de la tendencia del retardo, robusto frente al ruido de la señal de retardos, que permite obtener el plazo máximo esperado para que un paquete dado llegue al nodo.

4. Análisis de pérdidas y retardos con NRA

Dado que el algoritmo que se propone para la localización del mecanismo de recuperación se basa en el valor MOS estimado por E-model, es necesario obtener los parámetros requeridos en la versión simplificada (el retardo medio d de los paquetes y la probabilidad final de pérdidas e) para cada una de las alternativas.

Teniendo en cuenta el entorno de la Fig. 1, se asumirá que: el valor del retardo máximo permitido para una paquete d_{max} es de $300ms$, el periodo de generación de cada trama es de $t_f = 22ms$,

no se producen pérdidas en las solicitudes de retransmisión, y que el *jitter* experimentado es insignificante.

En el análisis del mecanismo activo de recuperación, denominaremos al procedimiento *NRA-DAS* cuando se utilice la detección *DAS*, y *NRA-DETAH* cuando se utilice la técnica de detección *DETAH*.

4.1. Retardos y probabilidad de pérdidas en NRA-DAS

Sea A_i el suceso de que un paquete perdido sea seguido de $i-1$ pérdidas, B_j el suceso de que un paquete tenga que ser retransmitido j veces para que llegue de S a NRA, y C_k el suceso de que un paquete deba ser retransmitido del nodo NRA al R k veces. La probabilidad $p(A_i \cap B_j \cap C_k)$ puede ser calculada para nuestro entorno como sigue en la expresión (12):

$$p(A_i \cap B_j) = \begin{cases} 1-p_1 & \text{si } i=0 \\ p_1^{i+j} \cdot (1-p_1)^2 & \text{si } i>0 \end{cases} \quad (10)$$

$$p(C_k) = (1-p_2) \cdot p_2^k \quad (11)$$

$$p(A_i \cap B_j \cap C_k) = p(A_i, B_j) \cdot p(C_k) \quad (12)$$

Para cada aparición del suceso $A_i \cap B_j \cap C_k$ existe un retardo asociado $r(A_i, B_j, C_k)$. La expresión correspondiente de retardo se muestra en la expresión (15):

$$r(A_i, B_j) = \begin{cases} i \cdot t_f + tp_1 \cdot (3 + 2 \cdot j) & \text{si } i \leq 0 \\ tp_1 & \text{si } i = 0 \end{cases} \quad (13)$$

$$r(C_k) = tp_2 \cdot (2 \cdot k + 1) \quad (14)$$

$$r(A_i, B_j, C_k) = r(A_i, B_j) + r(C_k) \quad (15)$$

Nótese en este caso que si $i=0$, no tiene sentido que j pueda tomar un valor distinto de 0.

En nuestro caso, la probabilidad extremo a extremo de llegada de un paquete tras los mecanismos de recuperación viene expresada por la ecuación (16), donde se calcula la suma de las probabilidades de que el paquete llegue antes del límite d_{max}:

$$p_{exito} = \sum_{i=1}^{i=i_{max}} \sum_{j=0}^{j=j_{max}} \sum_{k=0}^{k=k_{max}} p(A_i \cap B_j \cap C_k) + \sum_{k=0}^{k=k_{max}} p(A_0 \cap B_0 \cap C_k) \quad (16)$$

para valores de i, j y k tales que $r(A_i, B_j, C_k) < d_{max}$. Por tanto, la probabilidad de pérdida final como parámetro de entrada para el E-model se calculará como aparece en la expresión (17):

$$e_{NRA-DAS} = 1 - p_{exito} \quad (17)$$

De la expresión (15) podemos deducir los valores máximos i_{max}, j_{max} y k_{max} para cumplir $r(A_i, B_j, C_k) < d_{max}$, de forma que acotamos el número de operaciones para obtener las expresiones finales. Así, el retardo medio $d_{NRA-DAS}$ quedaría como aparece en la expresión (18):

$$d_{NRA-DAS} = \sum_{i=0}^{i=i_{max}} \sum_{j=0}^{j=j_{max}} \sum_{k=0}^{k=k_{max}} r(A_i, B_j, C_k) \cdot p(A_i \cap B_j \cap C_k) \quad (18)$$

4.2. Retardos y probabilidad de pérdidas en NRA-DETAH

Sea B_j el suceso de que un paquete tenga que ser retransmitido j veces para que llegue de S a NRA, y C_k el suceso de que un paquete deba ser retransmitido del nodo NRA al R k veces. La probabilidad $p(B_j \cap C_k)$ puede ser calculada para nuestro entorno como sigue en la expresión (19):

$$\begin{aligned} p(B_j) &= (1-p_1) \cdot p_1^j \\ p(C_k) &= (1-p_2) \cdot p_2^k \\ p(B_j \cap C_k) &= (1-p_1) \cdot p_1^j \cdot (1-p_2) \cdot p_2^k \end{aligned} \quad (19)$$

El retardo $r(B_j, C_k)$ asociado queda como sigue en la expresión (20):

$$\begin{aligned} r(B_j) &= p_1 \cdot (2 \cdot j + 1) \\ r(C_k) &= p_2 \cdot (2 \cdot j + 1) \\ r(B_j, C_k) &= r(B_j) + r(C_k) \end{aligned} \quad (20)$$

En nuestro caso, la probabilidad extremo a extremo de llegada de un paquete tras los mecanismos de recuperación viene expresada por la ecuación (21), donde se calcula la suma de las probabilidades de que el paquete llegue antes del límite d_{max}:

$$p_{exito} = \sum_{j=0}^{j=j_{max}} \sum_{k=0}^{k=k_{max}} p(B_j \cap C_k)$$

para valores de j y k tales que se cumpla la restricción $r(B_j, C_k) < d_{max}$. De (15) se pueden deducir los valores máximos j_{max} y k_{max}.

Por último, la probabilidad $e_{NRA-DETAH}$ de que expire el tiempo de validez de un paquete queda con (21):

$$e_{NRA-DETAH} = 1 - p_{exito} \quad (21)$$

El retardo medio $d_{NRA-DETAH}$ puede calcularse mediante la expresión (22):

$$d_{NRA-DETAH} = \sum_{j=0}^{j=j_{max}} \sum_{k=0}^{k=k_{max}} r(B_j, C_k) \cdot \cdot p(B_j \cap C_k) \quad (22)$$

5. Algoritmo de ubicación y selección del agente de recuperación

En el problema a resolver se presenta la misma topología descrita en la Fig. 1, con la salvedad de que existen varios nodos intermedios que podrían ejecutar el procedimiento NRA. Se trata pues de seleccionar el que vaya a resultar con el mejor valor MOS.

El algoritmo consiste, por tanto, en calcular en cada nodo intermedio con capacidad para ejecutar NRA el MOS a partir del E-model descrito, conociendo su retardo de propagación del emisor al nodo y del nodo al receptor (tp_1 y tp_2), así como la probabilidad de pérdida en ambos enlaces (p_1 y p_2). El cálculo se realiza para los distintos mecanismos de detección que incluya.

Cuando todos los nodos han calculado el valor MOS que les corresponde, se construye una lista ordenada por su puntuación MOS. De esta lista se escogerá de entre los primeras entradas cuyo valor MOS sea mayor o igual que $MOS_{max} - U_{MOS}$, donde MOS_{max} corresponde a la puntuación MOS más alta alcanzada, y U_{MOS} el umbral de error permitido. Dicho umbral debe tener un valor no superior a $U_{MOS} \leq 0.5$, ya que en ese caso se podría deteriorar demasiado la calidad que se obtendrá frente a la óptima para ese escenario.

La selección de entre los nodos y mecanismos permitidos de la lista se realizará de acuerdo a la disponibilidad de recursos en dicho nodo, así como la optimización de coste de migrar el agente de recuperación del nodo actual al nuevo seleccionado, u otros criterios de preferencia a la hora de ejecutar el procedimiento NRA en un nodo u otro.

El algoritmo se ejecutará en un nodo toda vez que haya una variación significativa de p_i o de tp_i.

6. Experimentación

Los experimentos realizados para comprobar la eficacia del algoritmo propuesto se realizan mediante simulación del escenario básico anteriormente descrito. En la ruta entre el emisor y el receptor se encuentran varios nodos con capacidad para ejecutar el mecanismo de recuperación descrito en el apartado 3.

Se generan distintos escenarios en los que se indica una probabilidad de pérdida extremo a extremo p y un retardo de propagación t_p desde el emisor al receptor. Además, se explicita un número de posibles nodos activos en la ruta de S a R. De esta forma, para cada escenario se generan aleatoriamente varias distribuciones de retardos de propagación entre nodos, y de probabilidades de pérdidas por enlaces, de modo que se cumplan las restricciones extremo a extremo.

Una vez generado un escenario, por cada nodo activo se calcula el valor de MOS esperado según las expresiones descritas en las secciones 2 y 4 para cada uno de los mecanismos propuestos NRA-DAS y NRA-DETAH.

Para cada escenario distinto (p, tp y número de nodos disponibles) se efectúa una simulación por cada posible combinación. Se esta forma se obtienen el MOS asociado al empleo del mecanismo de recuperación en cada uno de los nodos. Así se conoce la posición y versión del procedimiento de detección óptimas para ese escenario.

Por cada escenario se compararán esos valores con los obtenidos por el algoritmo propuesto. De esta forma se evaluará la eficacia del algoritmo, midiendo la diferencia o error entre el mayor valor MOS que se puede obtener y el valor MOS obtenido por desplegar el mecanismo de recuperación en el nodo seleccionado.

6.1. Parámetros de simulación

Se simulan 487 escenarios diferentes, creando distintos escenarios con retardos extremo a extremo $t_p = \{50ms, 20ms, 80ms, 100ms, 200ms\}$ y probabilidades de pérdida extremo a extremo p con valores en el rango $[0.1, 0.8]$, incrementados de 0.1 en 0.1. En cada uno de estos escenarios se realizan varias simulaciones en las que se introducen de 4 a 6 nodos activos intermedios, y se distribuyen de forma aleatoria.

Para medir la relevancia de un escenario en estas pruebas se escoge la distancia media del MOS esperado en cada nodo al mayor MOS alcanzable en el escenario. Cuanto mayor sea este valor, mayor será el impacto de seleccionar un nodo adecuado en la calidad final percibida. En las simulaciones realizadas, aproximadamente un

30 % de los escenarios tenían una distancia media de más de 0.5 puntos, diferencia significativa en esta escala.

En la Fig.2 se muestra la distribución acumulativa de distancia media de los MOS de todos los nodos al valor MOS óptimo para cada escenario. Como puede apreciarse, el 50 % de los escenarios tienen un valor de diferencia media de 0.25 puntos, y casi un 10 % un valor mayor a 1 punto en la escala MOS.

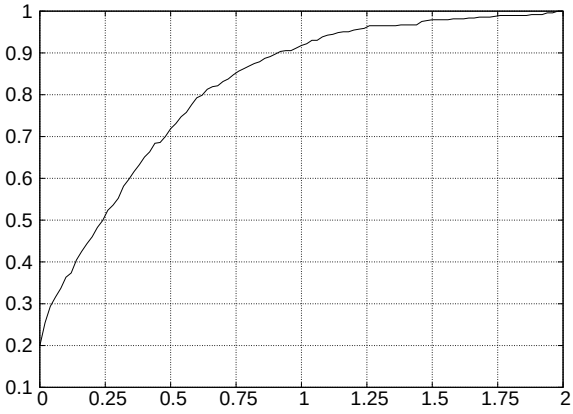


Figura 2: Distribución acumulada del valor medio de distancia al mayor MOS en los escenarios simulados.

6.2. Resultados de las simulaciones

Para caracterizar la eficacia del algoritmo de selección, se obtiene la diferencia entre el MOS obtenido en la simulación al ubicar el sevicio en el nodo seleccionado por el algoritmo propuesto, y el mejor MOS obtenido tras las simulaciones en todos los nodos posibles.

Los escenarios donde el empleo del algoritmo puede ser más relevante son aquellos en los que ubicar el servicio de recuperación en un nodo o en otro supongan una diferencia de MOS de más de 0.5 puntos. Por ello se escogen los escenarios de simulación donde la diferencia media de MOS de cada nodo frente al mejor que se pueda obtener sea mayor o igual a 0.5.

Así, la distribución de diferencias de puntuación o error en la escala MOS experimentados en los anteriores escenarios, se resume en la Fig. 3. Se puede observar cómo el 95 % de las elecciones del algoritmo obtienen un error inferior a 0.05 puntos en la escala MOS, y menos de un 1.5 % de las selecciones generan un error de entre 0.1 y 0.2 puntos en la escala MOS.

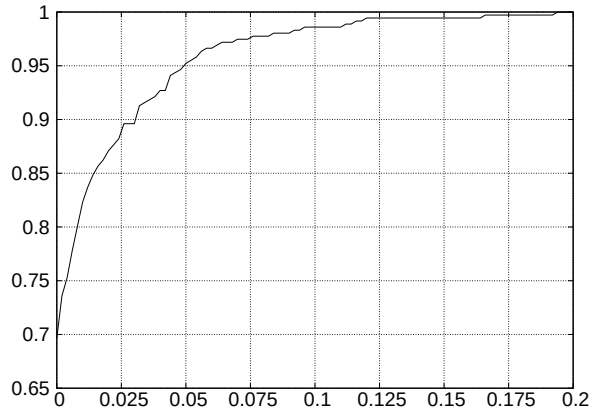


Figura 3: Distribución acumulada del error cometido tras la localización según el algoritmo.

7. Conclusiones

En el presente trabajo se diseña un algoritmo para la localización de un mecanismo de detección de pérdidas y retransmisión local, así como para la elección del mecanismo de detección asociado más indicado para las condiciones de la red en que se encuentra.

Dicho algoritmo se basa en la estimación de los retardos sufridos por los paquetes tras el proceso de recuperación, calculados a partir de la topología y las características de pérdidas del entorno en el que se ejecuta, teniendo como objetivo maximizar el valor perceptual que E-model proporciona.

Se observa tras la evaluación del algoritmo mediante simulación que el deterioro de la calidad percibida que introducen los errores de estimación del algoritmo en términos de la escala MOS no supera los 0.05 puntos para el 95 % de los casos.

Se demuestra, por tanto, que el algoritmo propuesto ofrece no sólo un procedimiento para elegir el mecanismo de detección en el nodo intermedio más adecuado, sino que además permite seleccionar de entre los mecanismos de detección que se dispongan el más ventajoso para el estado correspondiente de la red.

8. Trabajo futuro

Como trabajo futuro se planea ampliar el estudio incluyendo otros detectores, de forma que el algoritmo seleccione, de entre un abanico más amplio de posibilidades, el mecanismo de detección que mejor se ajuste al estado de la red.

Asímismo se pretende integrar el algoritmo de localización para su funcionamiento con otros

mecanismos de recuperación, como pueden ser técnicas de entremezclado, etc.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el proyecto de investigación TIC2002-02798 del Ministerio de Educación y Ciencia, financiado en un 70 % por fondos FEDER.

Referencias

[1] http://www.skype.com

[2] Juan J. Ramos Muñoz, Juan M. López Soler."Servicios de Multidifusión de Voz Sobre IP en Redes Activas. Mecanismos de Detección Activa de Pérdidas.". Monografía nº 55 del Dpto. de Electrónica y Tecnología de Computadores de la Universidad de Granada (2004). ISBN 84-7951-033-1

[3] Diane Kiwior, Stephen Zabele. "Active Resource Allocation in Active Networks ". IEEE Journal on Selected Areas in Communications, Vol 19 (3) (2001) 452-459.

[4] R.G. Cole, J.H. Rosenbluth. "Voice Over IP Performance Monitoring ".SIGCOMM Comput. Commun. Rev., vol 31 (2) (2001) 9-24.

[5] D.L. Tennenhouse, J.M. Smith, W.D. Sincoskie, D.J. Wetherall, G.J. Minden. "A Survey of Active Network Research". IEEE Communications Magazine 35 (1) (1997) 80-86.

[6] A. Banchs, W. Effelsberg, C. Tschudin, V. Turau."Multicasting Multimedia Streams with Active Networks". Proceedings of the 23rd Annual Conference on Local Computer Networks (LCN), IEEE, Lowell, MA, USA, (1998) 150-159.

[7] M. Calderón, M. Sedano, A. Azcorra, C. Alonse, "Active Network support for multicast applications", IEEE Network 12 (3) (1998) 46-52.

[8] M. Maimour, C. Pham," A Analysis of a router-based loss detection service for active reliable multicast protocols", en IEEE (Ed.), Proceedings of the IEEE International Conference On Networks (ICON 2002), IEEE, Singapoure, 2002.

[9] Juan J. Ramos Muñoz, Juan F. Núñez Negrillo, Juan M. López Soler. "Detección Activa de Pérdidas de Paquetes en Flujos de Audio en Tiempo Real". IV Jornadas de Ingeniería Telemática, Gran Canaria 2003.

[10] C. Holt. "Forecasting Seasonal and Trends by Exponentially Weighted Moving Averages". Research Memorandum 52, Office of Naval Research (1957).

Pago Anónimo en un Protocolo Verificable de Comercio Electrónico.

Magdalena Payeras Capellà, Josep Lluís Ferrer Gomila, Llorenç Huguet Rotger
Departament de Ciències Matemàtiques i Informàtica
Universitat de les Illes Balears
Carretera de Valldemossa, Km. 7,5. 07122, Palma de Mallorca (Illes Balears)
Teléfono: 971171390, Fax: 971173003
E-mail: mpayeras@uib.es

Abstract*. An electronic purchase represents an exchange between money and a digital product or the receipt of a physical product. Atomicity is a desired feature for electronic payment systems, allowing fair purchases. Some papers proposing fair payment have been published. In this paper we evaluate the role of the TTP in one of them, a fair exchange protocol useful in electronic purchases involving electronic coins, showing that the TTP can act maliciously, but in any case, with little modification to the original protocol, this behavior can be demonstrated, so the TTP is verifiable.*

1 Introducción

Algunos servicios electrónicos requieren el intercambio equitativo de elementos entre dos o más usuarios. El intercambio equitativo de valores siempre proporciona un tratamiento equitativo a todos los usuarios. Gracias a la equitatividad, al final de la ejecución de un intercambio, o todas las partes tienen el elemento que deseaban obtener, o en caso de que el intercambio no se haya finalizado, ninguna de las partes tiene el elemento deseado.

Entre las aplicaciones electrónicas que requieren un intercambio equitativo de información podemos encontrar la firma electrónica de contratos, el correo electrónico certificado y el pago a cambio de un recibo (o en el caso de la compra de productos digitales el intercambio de un pago directamente por un producto).

Una compra electrónica representa el intercambio de un pago por un recibo (o por un producto) en que el pago puede llevarse a cabo mediante diferentes tipos de sistemas. En la compra de un producto tangible, un recibo puede usarse como una prueba del pago para demostrar, sin posible rechazo por parte del comerciante de que el usuario ha realizado el pago. Cuando el pago se lleva a cabo en la compra de un producto digital, el intercambio del dinero por el producto puede llevarse a cabo directamente, pero se mantiene la necesidad de equitatividad en el intercambio, ya que el comprador no quiere tomar el riesgo de pagar sin la seguridad de que recibirá el producto, mientras el comerciante no quiere enviar el producto antes de recibir el pago, más aún teniendo en cuenta que los productos digitales pueden copiarse, y exigir el reintegro del producto, por consiguiente, no tiene sentido.

La clasificación de los protocolos de intercambio equitativo se basa en la presencia o ausencia de terceras partes de confianza (TTP) durante la ejecución del protocolo, y en su caso, de su grado de implicación.

- **Protocolos para intercambio equitativo que no requieren TTP.** Estos protocolos, por si solos, garantizan la seguridad del intercambio, y por consiguiente, no requieren la intervención de ninguna TTP. Esta independencia de una tercera parte es una característica deseable, pero algunos de estos protocolos requieren un gran número de interacciones entre las partes, así como una capacidad de cálculo elevada.
- **Protocolos para el intercambio equitativo de valores con la intervención de una TTP.** Entre ellos pueden distinguirse los protocolos que requieren la intervención de la TTP en cada ejecución del protocolo de aquellos dónde la TTP sólo interviene en caso de que el intercambio no concluya con éxito (estos últimos reciben el nombre de protocolos optimistas). La presencia constante de la TTP tiene sus consecuencias: el coste que el servicio representa para los usuarios, la posible congestión en la comunicación con la TTP y el retraso adicional causado por las comunicaciones entre el usuario y la TTP en cada ejecución del protocolo. En los protocolos optimistas la participación de la TTP se limita a algunos casos. Estos protocolos disponen de un subprotocolo de intercambio en el que la TTP no interviene. Si el subprotocolo de intercambio concluye satisfactoriamente, entonces la participación de la TTP no es necesaria en la ejecución del protocolo. Si no, se ejecutará otro subprotocolo que involucrará a la TTP. En función de las pruebas presentadas, la TTP podrá emitir los mensajes y tomar las decisiones que garanticen la equitatividad del intercambio.

De la clasificación anterior se deduce que por razones de eficacia es deseable que exista una TTP, pero que sólo intervenga para resolver disputas cuando la ejecución del protocolo lleve a una situación injusta. En consecuencia, las propiedades deseables se formulan para los intercambios optimistas [1]. Estas características son:

- **Equitatividad.** Cuando la ejecución del protocolo concluye, o todas las partes disponen de los objetos deseados, o ninguna de las partes dispone de ellos.
- **Asincronía.** Un protocolo proporciona asincronía si todas las partes pueden, en cualquier momento, concluir el intercambio conservando la equitatividad.
- **No repudio.** Después del intercambio, cada participante puede demostrar el origen del objeto que ha recibido, es decir, no puede negar la emisión del objeto propio.
- **Verificabilidad de la TTP.** Si la TTP interviene en la ejecución del intercambio y actúa incorrectamente, entonces la conducta fraudulenta debe ser demostrable.
- **Eficiencia.** Un protocolo eficiente utiliza el menor número posible de interacciones entre los usuarios.
- **Privacidad.** Un protocolo es confidencial si permite ocultar la información intercambiada, incluso a la TTP, si es el caso.

2 Pago por Producto o Recibo

En un pago mediante tarjeta de crédito, la orden de compra, que incluye el número de la tarjeta (y la firma) se intercambia por el recibo del pago o por el producto. El intercambio de una orden de compra firmada a cambio de un recibo del pago puede considerarse una aplicación de los protocolos de firma de contratos. Por otro lado, en los pagos con moneda electrónica, la moneda se convierte en el elemento a intercambiar por parte del comprador, y el intercambio no puede considerarse resuelto con los protocolos de firma de contratos, dado que existen situaciones específicas dónde la interrupción del intercambio podría causar la pérdida de la moneda para las dos partes o la pérdida de anonimato a alguna de ellas. Por ejemplo, cuando en un sistema off-line de pago mediante moneda electrónica un error causa que el pagador dude si el receptor ha recibido o no la moneda, el pagador no puede arriesgarse usando de nuevo la moneda, ya que si el pago hubiera concluido, el pagador no sólo seria identificado sino que también seria acusado de reutilización.

Además de proporcionar equitatividad, es deseable que los protocolos de intercambio permitan a las partes demostrar qué objeto han recibido de la otra parte, y por consiguiente, en caso de disputa posterior, poder presentar pruebas del intercambio. Los protocolos de compra pueden proporcionar intercambio atómico, entrega certificada para alguna o todas las partes involucradas en el intercambio, o ambas cosas.

En [10], la atomicidad del dinero se define como la característica que evita la creación o destrucción de dinero durante su transferencia. Por consiguiente, estos protocolos no proporcionan intercambio equitativo. La atomicidad de los bienes también se define en [10] y se aplica a los protocolos que no sólo presentan atomicidad del dinero sino que también permiten el intercambio equitativo entre el producto y la moneda.

La entrega certificada [10] proporciona atomicidad del bien y de la moneda, y también proporciona a ambas partes pruebas sobre lo que han enviado y sobre lo que la otra parte ha recibido. La entrega certificada puede ser unilateral o bilateral [4], en función de cuántas partes posean pruebas de recepción. La entrega certificada y atómica [8] proporciona atomicidad del bien y de la moneda y las partes se ponen de acuerdo en la negociación inicial. El intercambio proporciona pruebas de que tanto el bien como la moneda se han recibido. Presenta, al mismo tiempo, atomicidad del bien y entrega certificada. Finalmente, la compra atómica distribuida [8] proporciona atomicidad del dinero y del bien cuando más de un comerciante está implicado en la compra.

La solución adoptada en [10] es útil en caso de problemas de comunicación, pero no es útil en caso del intento de fraude. El sistema usa un coordinador que conoce la identidad de todas las partes, por lo que el sistema no permite pagos anónimos. Los protocolos descritos en [4, 6, 9, 10] llevan a cabo el intercambio con una TTP en línea. En [6], la TTP activa es un *board* dónde todos los usuarios pueden leer y escribir. [4] proporciona entrega certificada unilateral, y el banco, que actúa como TTP, está implicado en el pago. [10] utiliza pago on-line en el que la TTP actúa también como banco y garantiza el intercambio equitativo durante el pago. [8] y [9] son soluciones similares donde se utiliza un coordinador de pago on-line.

Otras soluciones, como [5], no necesitan una TTP. En este caso se opta por dividir la moneda en dos partes que se enviarán antes y después de la recepción del bien. El comerciante no esta protegido; no puede contactar con una TTP si no recibe la segunda parte de la moneda. Las monedas pueden tener un estado ambiguo si el comprador no se arriesga a ser identificado en caso de reutilización. Como conclusión, no proporciona atomicidad, y sólo proporciona una pequeña protección al pagador. [13] no satisface las características ideales, si el intercambio no finaliza de forma satisfactoria, el cliente no podrá conseguir el bien, sólo podrá recuperar el dinero, es decir, el intercambio puede ser cancelado, pero no finalizado. [12] no incluye el análisis del sistema del pago que se usaría en el intercambio. La compra no es certificada; el comerciante no puede demostrar que el cliente ha recibido el bien.

Según las características ideales, el objetivo es una compra certificada y atómica en un protocolo que proporcione anonimato, por lo menos al pagador, manteniendo el anonimato que proporciona el sistema del pago y con la posibilidad de verificar el comportamiento de la TTP.

A continuación resumimos nuestro protocolo original, publicado en [7], el cual va a ser analizado en la sección siguiente y posteriormente mejorado en la sección 5. El protocolo presenta las características siguientes:

- **Entrega certificada bilateral:** El comerciante puede demostrar que el comprador ha recibido el producto o recibo. Por otra parte el comprador puede demostrar que el comerciante ha recibido el pago, así como qué elemento ha recibido él.
- **Anonimato:** El comprador será anónimo si el sistema del pago usado en el intercambio es uno anónimo, y permanecerá anónimo aunque contacte con la TTP. Si el intercambio concluye con la participación de la TTP y el cliente usa la moneda en otro establecimiento, la reutilización se descubrirá como en el caso habitual y el comprador será identificado.
- **Intercambio:** en el proceso de intercambio, el pago se lleva a cabo en dos fases. En la primera una moneda se envía al comerciante mientras en la segunda se envía una prueba confidencial relacionada con la moneda, sólo conocida por el pagador. El receptor no puede depositar la moneda si no recibe la segunda parte del pago. Sin embargo, con la primera parte de la moneda el receptor puede contactar con la TTP para finalizar el intercambio, pero no para depositar la moneda.
- **Seguridad del pago:** El intercambio mantiene las características de seguridad del sistema de pago utilizado: descubre la reutilización, identifica a los reutilizadores y previene la sobreutilización, el robo y la falsificación.
- **TTP off-line:** La TTP sólo participa en el proceso de resolución de conflictos cuando el subprotocolo de intercambio no se ha completado o alguna parte ha actuado fraudulentamente.
- **Eficaz y funcional con los sistemas del pago habituales:** El protocolo de intercambio es apropiado para su uso con varios sistemas de pago mediante moneda electrónica. Las características que estos sistemas deben satisfacer son:
  - Moneda creada por el banco (sistema de débito).
  - El banco no puede relacionar la moneda con la identidad del pagador: monedas anónimas.
  - El comerciante puede verificar la moneda cuando la recibe. No puede prevenir la reutilización.
  - El pago tiene una fase de reto-respuesta.
  - Se identifica a los reutilizadores a posteriori.
  - El pagador permanece anónimo si se comporta correctamente.

  Estas características se dan en numerosos protocolos del moneda electrónica, entre ellos [2] y [3], sistemas que han sido adaptados para demostrar la validez del protocolo del intercambio.
- **Finalización del intercambio:** una vez el compromiso de la compra ha sido establecido (2 pasos), el protocolo permite finalizar el intercambio, no sólo cancelarlo.

3. Descripción del protocolo

Tres partes participan en el protocolo de intercambio: el comprador (o pagador), el comerciante (o receptor) y la TTP. El comprador quiere comprar un producto identificado como *Código_producto*. La notación usada en la descripción del protocolo se incluye en la tabla siguiente:

Tabla 1: Notación

C	Comprador anónimo
M	Receptor
T	TTP
H[]	Función de hash
E_k[]	Cifrado con la clave secreta *k*
D_k[]	Descifrado con la clave secreta *k*
PR_x[]	Cifrado con la clave privada de *x*
PU_x[]	Cifrado con la clave pública de *x*
Sign[x,y]	Firma sobre *x* que prueba el conocimiento de un elemento secreto, *y*
α	Elemento secreto
Código_producto	H[descripción_producto]
CANCELADO, FINALIZADO	Variables booleanas. Valor por defecto: Falso

El protocolo original, descrito en [7] se ha modificado para obtener verificabilidad de la tercera parte de confianza. El protocolo esta formado por un subprotocolo de intercambio y dos subprotocolos adicionales.

El subprotocolo de intercambio contiene los siguientes pasos:

- **Paso 0. Selección del producto y orden de compra.** *C* envía la orden de compra que contiene *Código_producto* al comerciante. *C* envía la moneda que se utilizará en el pago a *M*.
- **Paso 1. Primera parte del compromiso de compra: reto.** *M* genera un reto para el pago (*rp*). Éste será el reto usado en el sistema de pago electrónico. *M* cifra el producto o recibo usando la clave de sesión *k*. Esta clave se cifra con la clave pública de *T*. Finalmente, *M* firma la relación entre los elementos y los envía a *C*.
- **Paso 2. Segunda parte del compromiso de compra: respuesta al reto.** *C* responde al reto del pago y firma la relación entre la moneda y el producto o recibo cifrado (*c*), demostrando que conoce la segunda parte de la moneda (elemento secreto α). La respuesta al reto del pago, *rrp*, puede usarse para identificar al cliente en caso de reutilización. Una vez recibido este mensaje, ambas partes pueden pedir la finalización del intercambio.
- **Paso 3. *M* envía la clave de sesión.** Después de la recepción de paso 2, *M* verifica las respuestas recibidas de *C*: *Sign(d, α)* y *rrp* y envía la clave *k*, necesaria para descifrar el producto o recibo.
- **Paso 4. *C* envía la prueba confidencial**. *C* envía la prueba confidencial que permitirá el depósito de la moneda.

Tabla 2. Subprotocolo de intercambio

SUBPROTOCOLO DE INTERCAMBIO	
0. C → M:	Código_producto, Moneda
1. M → C:	rp, c = E_k(producto), $K_t = PU_T(k)$, $H_M = PR_M${H[H(c), Kt], Id}
2. C → M:	rrp, d = H[Moneda, c, Id], Sign(d, α)
3. M → C:	$K_M = PR_M$ (k, Id)
4. C → M:	α

Los pasos 1 y 2 del subprotocolo de intercambio forman el compromiso de compra. Después del paso 2, *T* puede finalizar el intercambio a petición de *C* o *M* (subprotocolo de finalización). Si el intercambio se detiene antes de la recepción de paso 2, el compromiso no se establece, y *T* no puede concluirlo. Para invalidar los elementos enviados en el paso 1, *M* puede pedir la cancelación del intercambio mediante el subprotocolo de cancelación. El cuarto paso del subprotocolo de intercambio es necesario; sin él el receptor de la moneda, después del paso 2, podría actuar maliciosamente y solicitar la cancelación alegando no haber recibido la moneda. Entonces, si *C* usara la moneda de nuevo perdería el anonimato, y su identidad sería revelada.

Los subprotocolos de cancelación y de finalización se ejecutan entre *C* o *M* y *T*, cuando el subprotocolo de intercambio no ha concluido con éxito. *T* puede escoger entre concluir o cancelar el intercambio en función de las pruebas presentadas y del orden de compra.

El subprotocolo de cancelación solo puede ser ejecutado por *M* en caso de que el compromiso de compra no concluya. El subprotocolo de finalización puede ser ejecutado por ambas partes una vez el compromiso de compra ha concluido, es decir, si *C* no recibe la clave *k* (paso 3) o si el comerciante no recibe la prueba secreta de la moneda, α (paso 4).

Tabla 3. Subprotocolo de cancelación

SUBPROTOCOLO DE CANCELACIÓN		
	M → T:	Moneda, c, k_T, h_M, $h_{MT1} = PR_M$(c, k_t, h_m, moneda)
IF (FINALIZADO = TRUE)	T → M:	rrp, d, Sign(d, α), $P_{TM} = PR_T(\alpha)$
ELSE	T → M: T:	Prueba de cancelación = PR_T("cancelado", h_m) CANCELADO = TRUE

Tabla 4. Subprotocolo de Finalización de C

SUBPROTOCOLO DE FINALIZACIÓN DE C		
	C → T:	Moneda, rp, c, k_T, h_M, rrp, d, Sign(d, α), α
IF (CANCELADO = TRUE)	T → C:	Prueba de Cancelación = PR_T("cancelado", Sign (d, α))
ELSE	T → C: T:	$PR_T(k)$ FINALIZADO = TRUE

Tabla 5. Subprotocolo de Finalización de M

SUBPROTOCOLO DE FINALIZACIÓN DE M		
	M → T:	Moneda, rp, c, k_T, h_M, rrp, d, Sign(d, α), h_{MT2} = PR_M(c, k_t, h_m, moneda, rrp, Sign (d, α))
IF (FINALIZADO = TRUE)	T → M:	$P_{TM} = PR_T(\alpha)$
ELSE	T → M: T:	Autorización de depósito sin α FINALIZADO = TRUE CANCELADO = FALSE

4. Equitatividad

Para evaluar si el protocolo es equitativo, analizaremos todas las posibles situaciones derivadas de la ejecución del protocolo, involucrando o no a la TTP.

- **Intercambio concluido.** Si el intercambio se ha llevado a cabo sin problemas, *C* dispone del producto o recibo ($D_k(c)$) y puede demostrar cual es el producto recibido ($H_M = PR_M\{H[H(c), K_i], Id\}$). Es más, *C* puede demostrar que llevó a cabo el pago, ya que dispone de la clave: $K_M = PR_M(k, Id)$. *M* tiene ambas partes del pago: la moneda, *rrp*, *Sign(d, α)* y la prueba secreta de la moneda: α. Con este elemento puede demostrar que *C* ha recibido el producto o recibo.
- **Intercambio inacabado.** Si el intercambio no concluye con éxito, ambas partes pueden contactar con *T* y solicitar la finalización o la cancelación del intercambio. El intercambio puede haberse interrumpido en diferentes puntos:
  - ***M no recibe el mensaje de paso 2.***

    Si el paso 1 o el paso 2 no se completan, el compromiso de compra no está establecido. *M* puede pedir la cancelación del intercambio mientras *C* puede pedir su finalización. *M* no puede pedir la finalización ya que no dispone del elemento *Sign(d, α)*.
    1. ***C* finaliza, *M* cancela:** *T* envía la clave *k* a *C* y α a *M.*
    2. ***M cancela, C* finaliza*:*** *T* envía una prueba de cancelación a *M. C* no recibirá la clave*, k.*

  - ***C* no recibe el mensaje del paso 3 o *M* no recibe el mensaje del paso *4.***

    M puede finalizar o cancelar el intercambio mientras *C* sólo puede finalizar el intercambio.
    1. ***M* finaliza, *C* finaliza*:*** *M* obtendrá una autorización para depositar sin α. Cuando *C* intente finalizar, *T* le enviará la clave.
    2. ***M* cancela, *C* finaliza*:*** *M* y *C* obtendrán una prueba de cancelación.
    3. ***C* finaliza, *M* cancela o *C* finaliza, *M* finaliza:** *C* obtendrá la clave *k* y *M* obtendrá α.

En todo caso, la ejecución de los subprotocolos conduce a una situación equitativa.

5. Verificabilidad de la Tercera Parte de Confianza.

Durante la ejecución de los subprotocolos de cancelación o finalización, *T* decide el estado final del intercambio a partir de los valores de dos variables booleanas y de la información recibida en la solicitud. Si *T* no sigue el subprotocolo y envía los elementos inadecuados a las partes, estará actuando incorrectamente. Las partes, o un verificador externo, han de poder descubrir y demostrar la conducta fraudulenta de *T*. Si el fraude puede descubrirse y demostrarse, el protocolo será verificable.

Todas las posibles conductas fraudulentas por parte de *T* se listan a continuación:

- Si *M* no recibe el mensaje de paso 2, los usuarios pueden actuar de la siguiente forma:
  - ***M* cancela, *C* finaliza.** En este caso *T* puede actuar incorrectamente, dando una prueba de cancelación a *M*, y revelando la clave *k* a *C*. Esta conducta fraudulenta se llamará **FB1**. *T* no puede enviar α a *M,* ya que ignora su valor.
  - ***C* finaliza, *M* cancela.** Si *T* no proporciona la clave *k* a *C*, alegando que el intercambio ya ha sido cancelado, y, por otra parte, envía α a *M*, estará de nuevo cometiendo un fraude. Esta situación se llamará **FB2**. En este mismo caso, *T* puede proporcionar la clave *k* a *C* aunque no envíe α a *M*, alegando que el intercambio no ha concluido. Esta es la conducta anteriormente descrita como **FB1**.
- Si *C* no recibe el mensaje del paso 3 o *M* no recibe el mensaje del paso 4, las situaciones siguientes son posibles:
  - ***M* finaliza, *C* finaliza.** *T* no envía *k* a *C,* y autoriza a que *M* deposite la moneda sin el conocimiento de la prueba secreta, α. Esta conducta se llamará **FB3**.
  - ***C* concluye, *M* finaliza. FB3** es de nuevo posible**.** Es más, *T* puede proporcionar la clave *k* a *C* y puede autorizar a *M* el deposito de la moneda sin el conocimiento de α. Esta conducta se llamará **FB4**.

Se han descrito cuatro conductas fraudulentas diferentes. A continuación se demostrará que todas pueden descubrirse, así como demostrar el comportamiento incorrecto de *T*.

• **FB1**: *M* tiene una prueba de la cancelación y *C* tiene la clave, *k*. Esta situación puede ser el resultado de dos tipos diferentes de ejecución. En la primera, *C* obtiene la clave de *M*. Después, *M* pide la cancelación del intercambio y obtiene una prueba de cancelación. *T* ha actuado correctamente, es *M* quien ha actuado fraudulentamente, ya que pudo solicitar la finalización del intercambio. *C* puede demostrar el fraude de *M* mostrando la firma de *M* sobre *k,* enviada en el paso 3 del subprotocolo de intercambio, cuando *T* dispone de h_{MTI}. La segunda situación se produce cuando *M* solicita la cancelación del intercambio y *T* envía la clave *k* a *C*. *M* puede solicitar a *C* que muestre su firma sobre *k* para demostrar la conducta fraudulenta de *T*. Si *C* no puede proporcionarlo, la conducta fraudulenta de *T* puede demostrarse (*C* tiene $PR_T(k)$).

• **FB2**: Una vez *C* haya obtenido una prueba de cancelación, puede contactar con el banco para depositar la moneda o pedir su cambio por una moneda no usada en un intento de compra, sin riesgo alguno. En este momento, el banco descubre que el dinero ya ha sido depositado por parte de *M*, y por consiguiente, al proporcionar una prueba de cancelación a *C*, *T* actuó incorrectamente, a menos que *T* disponga de H_{MTI}, ya que en este caso *T* demuestra que fue *M* quien actuó incorrectamente cuando solicitó la cancelación del intercambio.
• **FB3**: Este caso se demuestra de la misma forma que FB2
• **FB4**: M puede depositar la moneda, pero también puede demostrar que, aunque sin perjuicio alguno, *T* ha actuado incorrectamente, enviando *k* a *C* si no dispone de K_M y por otro lado dispone de $PR_T(k)$, ya que *T* debería haber enviado α a *M*, en lugar proporcionarle una autorización para depositar la moneda.

En cualquier caso, la conducta incorrecta de *T* puede demostrarse, y por consiguiente la *TTP* es verificable.

6. Conclusiones

El intercambio equitativo dónde uno de los valores a intercambiar es una moneda electrónica tiene lugar en la compra electrónica de un producto. El valor a intercambiar será el producto o un recibo, dependiendo del tipo de producto (digital o físico). En [7] presetaos un protocolo de intercambio equitativo para sistemas de pago existentes en que el comprador y el comerciante pueden intercambiar sus elementos con sólo 4 pasos sin la intervención de la TTP. En este protocolo, sin embargo, la TTP puede ser invocada para la resolución de disputas. Por ésta razón, una conducta fraudulenta de la tercera parte puede provocar que el intercambio deje de ser equitativo. En este artículo explicamos cómo el protocolo puede modificarse para permitir el descubrimiento de los intentos de fraude por parte de la TTP. De esta forma la TTP es verificable y el intercambio siempre finaliza de forma equitativa.

Referencias

[1]. Asokan, N., Shoup, V., Waidner, M.: "Asynchronous protocols for optimistic fair exchange", IEEE Symposium on Research in Security and Privacy, páginas 86-99, 1998.

[2]. Brands, S.: "Untraceable off-line cash in wallet with observers", Crypto'93, LNCS 773, páginas 302-318, Springer Verlag, 1994.

[3]. Chaum, D., Fiat, A., Naor, M.: "Untraceable electronic cash", Crypto'88, LNCS 403, páginas 319-327. Springer Verlag, 1988.

[4]. Camp, J.,Harkavy, M., Tygar, J.D., Yee, B.: "Anonymous atomic transactions", 2nd USENIX workshop on electronic commerce, páginas 123-133, 1996.

[5]. Jakobsson, M.: "Ripping coins for a fair exchange", Eurocrypt'95, LNCS 921, páginas 220-230, Springer Verlag, 1995.

[6]. Pagnia, H., Jansen, R.: "Towards multiple payment schemes for digital money", Financial Cryptography' 97, LNCS 1318, páginas 203-216, Springer Verlag, 1997.

[7]. Payeras, M., Ferrer, J. L., Huguet, L.: "Incorporación de Atomicidad a los Protocolos de Pago Electrónico: Intercambio Equitativo de Moneda Electrónica por Producto o Recibo", Novática, N. 170, páginas 57-60, Julio-Agosto 2004.

[8]. Schuldt, H., Popovivi, A., Schek, H.: "Execution guarantees in electronic commerce payments.", 8th international workshop on foundations of models and languages for data and objects (TDD'99), LNCS 1773, Springer Verlag, 1999.

[9]. Su, J., Tygar, J.D.: "Building blocs for atomicity in electronic commerce", 6th USENIX security symposium, 1996.

[10]. Tang, L.: "Verifiable transaction atomicity for electronic payment protocols", IEEE ICDCS'96, páginas 261-269, 1996.

[11]. Tygar, J.D.: "Atomicity in electronic commerce", 15th ACM symposium on distributed computing", páginas 8-26, 1996.

[12]. Vogt, H., Pagnia, H., Gärtner, F.C.: "Modular fair exchange protocols for electronic commerce" 15th Annual Computer Security Applications Conference,'99, páginas 3-11, 1999.

[13]. Xu, S., Yung, M., Zhang, G., Zhu, H. "Money conservation via atomicity in fair off-line e-cash", International security workshop ISW'99, LNCS 1729, páginas 14-31, Springer Verlag, 1999.

Análisis de Seguridad en Redes Inalámbricas de Sensores

Rodrigo Román Castro[1], Javier López Muñoz[1], Jianying Zhou[2]
[1] E.T.S. Ingeniería Informática, Universidad de Málaga, 29071 - Málaga
[2] Institute for Infocomm Research, 21 Heng Mui Keng Terrace, Singapore 119613
roman@lcc.uma.es, jlm@lcc.uma.es, jyzhou@i2r.a-star.edu.sg

Abstract *The design and development of security infrastructures and protocols for Wireless Sensor Networks is a difficult task, due to several factors like the constraints of the sensor nodes and the public nature of the communication channels. The intrinsic features of these networks create numerous security problems. In this paper, we analyze and put into perspective those problems.*

1. Introducción

Las redes inalámbricas de sensores (Wireless Sensor Networks) [1] están compuestas por cientos o miles de dispositivos (nodos) equipados con sensores (temperatura, sonido, movimiento,...) y con una capacidad limitada de comunicación y cálculo. Estas redes permiten a los sistemas informáticos acceder y procesar a distancia información procedente del mundo real.

Los campos de aplicación de las redes de sensores son muy variados (Salud, Entornos Inteligentes,...), y están creciendo día a día. Actualmente, los usos más comunes de este tipo de redes incluyen la monitorización de espacios naturales, la seguridad en la construcción y mantenimiento de edificios, la vigilancia de espacios protegidos, y el control de maquinaria industrial.

Sin embargo, las redes de sensores son extremadamente vulnerables ante cualquier tipo de ataque, tanto interno como externo. Esto es debido a factores tales como las limitaciones de los nodos, su falta de protección ante ataques físicos, y la facilidad de acceso al canal de comunicaciones. En este contexto, todo protocolo, arquitectura, o aplicación que no tome en cuenta la seguridad desde las primeras fases de su desarrollo difícilmente podrá ser utilizado en la vida real.

El objetivo de este artículo es el de analizar los problemas de seguridad en redes de sensores, dando además una perspectiva de algunas de las soluciones que actualmente se pueden aplicar, pero que en muchas ocasiones no son las óptimas. La estructura del artículo es la siguiente: En la sección 2, se muestra la infraestructura y los elementos que componen una red de sensores inalámbrica. La sección 3, que es el núcleo de este trabajo, analiza los distintos problemas de seguridad asociados con este tipo de redes, tales como el uso de primitivas de seguridad, la infraestructura de claves, el encaminamiento de información, etc. Finalmente, la sección 4 finaliza el artículo discutiendo los principales retos de seguridad actuales.

2. Infraestructura de Redes de Sensores

La infraestructura de una red de sensores se divide en dos partes, la *red de adquisición de datos* y *la red de diseminación de datos.*

- La red de adquisición de datos es la Red de Sensores propiamente dicha. Esta formada por un conjunto de nodos cuya tarea es la de medir, procesar y reenviar los datos físicos de sus alrededores, y por una o más estaciones base (Base Station) a cargo de recoger los datos procedentes de los nodos y de enviarles información de control procedente de los usuarios.

- La red de diseminación de datos es un conjunto de redes, tanto inalámbricas como basadas en cables, que proporcionan a cualquier usuario una interfaz con la que interactuar con la red de sensores. La seguridad de esta red esta fuera del alcance de este artículo.

Los nodos están densamente distribuidos o muy cerca o en el interior del objeto u entorno que desea observarse, y las medidas realizadas deben enviarse hacia la estación base donde los usuarios puedan acceder a ellas. Todos estos nodos son dispositivos con unos recursos (memoria, capacidad computacional, batería) muy limitados. Por otro lado, las estaciones base no poseen tantas restricciones como los nodos, y suelen tener un suministro continuado de energía.

Actualmente, el modelo de nodo más popular es el MICA2 [2]. Éste utiliza un procesador de 8 Mhz con 128Kb de memoria de instrucciones, 4Kb de RAM, y 512Kb de memoria flash para almacenamiento de datos. Su transmisor de radio le permite enviar 19.2Kb/s en un canal compartido, y su batería le permite trabajar de forma ininterrumpida hasta 2 semanas, aunque es posible mantenerlo en funcionamiento durante 1 año.

Debido a las limitaciones de su infraestructura, una red de sensores es extremadamente vulnerable ante cualquier tipo de ataque, tanto procedente del exterior (inyección de paquetes) como del interior (un nodo controlado por un adversario). Por lo tanto, es necesario que tanto la infraestructura como los protocolos de la red estén preparados para afrontar este tipo de situaciones adversas. Proteger la información no solo requiere de un conjunto de algoritmos eficientes de cifrado, sino de una política óptima de manejo de claves en términos de distribución, almacenamiento y mantenimiento. Además, es necesario proteger la agregación de datos dentro un grupo (estático o dinámico) de nodos y su encaminamiento hacia la estación base. Finalmente, entre otras cosas, la red debería ser capaz de monitorizar errores o brechas de seguridad en cualquiera de sus miembros y responder ante estas circunstancias de forma automática.

3. Seguridad y Redes de Sensores

3.1. Primitivas de Seguridad

Los nodos que forman parte de una red de sensores utilizan transmisores de radio para sus comunicaciones. Todos los nodos existentes en el mercado operan en bandas de frecuencia que no necesitan de licencia, sean los 433 Mhz (el espectro más bajo de las bandas ISM en Europa) o las bandas utilizadas en el estándar IEEE 802.15.4 para redes de área personal (PAN) [3]. La capacidad máxima del canal de comunicación oscila entre 19.2 Kbps y 250 Kbps.

Cualquier adversario puede acceder a la información procedente de una red de sensores, debido a que los nodos están (normalmente) distribuidos en un entorno de fácil acceso, y los canales de comunicación inalámbricos son inherentemente inseguros. En consecuencia, cualquier dispositivo puede escuchar o inyectar paquetes en la red de sensores.

Es por lo tanto indispensable incluir unas primitivas de seguridad dentro de los nodos para así proporcionar tanto una mínima protección al flujo de información como una base para la creación de protocolos seguros. Esas primitivas de seguridad son la *criptografía de clave simétrica* (SKE), los *códigos de autenticación de mensajes* (MAC), y la *criptografía de clave pública* (PKC). Debido a la escasez de recursos disponibles en los nodos, implementar estas primitivas de seguridad de una forma eficiente (usando menos energía, memoria y ciclos de CPU) sin sacrificar sus propiedades de seguridad es todo un reto.

Los nodos comerciales disponibles actualmente son capaces de implementar SKE a nivel software de una forma eficiente en términos de CPU, memoria y energía. Un ejemplo es el proyecto TinySec [4], librería criptográfica incluida dentro del sistema operativo TinyOS. TinySec es capaz tanto de autenticar y verificar la integridad de un mensaje como de proteger su confidencialidad, o ambas, utilizando cifrados de bloque como Skipjack o RC5 en modo CBC. En todos los casos el gasto de energía, ancho de banda y CPU es menor del 10 %.

En nodos cuya radio funcione de acuerdo al estándar 802.15.4 [3] la SKE la proporciona el hardware, quitando carga a la CPU y disminuyendo el uso de energía del nodo. En éste estándar, una aplicación puede elegir entre varias "suites" de seguridad que proporcionan (juntas o por separado) cifrado, autenticación y protección contra el reenvío utilizando el algoritmo AES. No obstante, no todos estos modos de funcionamiento son seguros [5], por lo que los diseñadores de aplicaciones deben tener cuidado al usar el estándar.

Respecto al MAC, éste se suele calcular utilizando algoritmos de cifrado en bloque, en un modo especial denominado CBC-MAC. Este modo es eficiente y rápido, y además permite reducir la cantidad de memoria requerida para implementar el MAC al compartir el algoritmo de cifrado en bloque con los demás módulos criptográficos del nodo, sean éstos software [4] o hardware [3].

En el contexto de las redes de sensores, la inclusión de PKC en un nodo utilizando software se consideraba imposible, pero no existían experimentos que demostraran esa presunción. Algunos estudios apuntaron a la criptografía de curva elíptica (ECC) como una posible solución aplicable a las redes de sensores, debido al reducido tamaño de sus claves, la rapidez de cálculo de sus primitivas, y los ahorros en energía y memoria en comparación con otros algoritmos como RSA.

Finalmente, un trabajo reciente en este área [6] desarrolló una implementación de PKC en TinyOS. Ésta utiliza ECC sobre $\mathbb{F}_{2^p}$ con una longitud de clave de 163 bits, con un gasto de memoria de 1Kb de RAM y 34Kb de ROM, y con un tiempo de ejecución de 34 segundos tanto para la generación de claves como para la generación de una clave secreta compartida.

3.2. Infraestructura de Claves

Los canales de comunicación entre dos nodos cualesquiera de la red de sensores deben estar protegidos para evitar ataques procedentes de agentes externos a la red. Esta protección la proporcionan las primitivas de seguridad introducidas en la sección anterior, pero para su uso es necesario que cada nodo pueda disponer de una serie de claves. Es por tanto necesario desarrollar una infraestructura de claves.

Existen tres factores básicos en el diseño de una infraestructura de claves para redes de sensores: almacenamiento, distribución, y mantenimiento de claves.

- Las políticas de almacenamiento indican el número de claves que un nodo necesita almacenar para abrir un canal de comunicación seguro con otros miembros de la red. Influye sobre la solidez de la red (network resilience), que define el porcentaje de la red que puede ser controlado por un adversario después de que éste obtenga las claves de un subconjunto de los nodos, y también influye sobre la cantidad de memoria disponible para el nodo.

- Los protocolos de distribución definen como se distribuyen las claves a los diversos nodos. Un nodo puede recibir sus claves antes de incorporarse en la red de sensores, o crear sus claves después (dentro de la red) utilizando información previamente almacenada.

- Los protocolos de mantenimiento especifican como un nodo puede incluirse o eliminarse de la red de sensores, recibiendo una serie de claves o anulando el uso de las que ya disponía. Este área de la infraestructura de claves no se encuentra muy desarrollada.

Respecto al almacenamiento de claves, existen dos casos extremos de diseño: *modo de clave global* (global keying) y *modo de clave por parejas* (pairwise keying). En el modo de clave global, existe una sola clave que todos los nodos poseen e utilizan para cifrar sus canales de comunicación. En el otro modo, clave por parejas, un nodo debe almacenar una clave por cada uno de los otros nodos existentes en la red, de tal forma que cada par de nodos compartirá un canal seguro específico.

Ninguno de los casos anteriores es viable en la mayoría de los escenarios posibles. El modo de clave global no proporciona solidez a la red, ya que si un solo nodo revela su clave a un adversario, todas las comunicaciones de la red se verán comprometidas. Y el modo de clave por parejas no es una solución escalable, debido a las restricciones de memoria de los nodos. Por esta razón se han estado buscando soluciones más óptimas, tales como compartir claves únicamente entre vecinos, o el paradigma de los conjuntos de claves (key pools).

El paradigma de los conjuntos de claves, introducido en [7], busca obtener un equilibrio entre el número de claves distribuidas en cada nodo y la solidez de la red. En este paradigma todos los nodos recogen un número determinado de claves de un conjunto global, creando conjuntos locales, antes de ser incluidos en la red de sensores. Después, solo los nodos que compartan una clave (o un número determinado de claves) de sus propios conjuntos pueden abrir un canal seguro de comunicación. El tamaño del conjunto global y de cada conjunto local son factores que influyen en la memoria disponible de los nodos y en la conectividad y la solidez de la red.

Este paradigma ha sido mejorado posteriormente, buscando optimizar o la construcción del conjunto global o la distribución de las claves hacia los conjuntos locales, de tal forma que la conectividad de la red sea cercana al 100 % (cada nodo pueda comunicarse con su vecino directo) mientras se disminuye el tamaño de los conjuntos locales y se aumenta la solidez de la red. Existen varias soluciones que logran este objetivo, sean basadas en principios matemáticos (como el esquema de Bloom [8] o la teoría combinatoria [9]), o aprovechando información obtenida "a priori" respecto a la distribución física de los nodos en la red de sensores [10].

Otros protocolos son capaces de negociar las claves de un nodo una vez que éste se haya sido incorporado a la red de sensores. En una de las soluciones un nodo negocia las claves que compartirá con sus vecinos más directos a través de la estación base [11], aunque este método puede no ser escalable. En otro modelo más simple, cada nodo contacta con sus vecinos y negocia las claves justo después de la construcción de la red [12]. En este modelo no se protege el intercambio de información, ya que en la mayoría de los escenarios no existe ninguna amenaza en el momento de la creación de la red de sensores.

Un área que aún esta inexplorada es el uso de criptografía de clave pública para la negociación de claves entre pares de nodos. Ya que es posible utilizar PKC en redes de sensores [6], queda por investigar como aplicarla en la creación e intercambio de claves y en los protocolos de mantenimiento de claves.

3.3. Infraestructura de Clave Local - Grupos Seguros

A lo largo de la vida útil de una red de sensores, existen ciertas situaciones en las que uno o más subconjuntos de nodos deben agruparse para cooperar en una tarea determinada. Un ejemplo de esta cooperación es cuando un grupo de nodos recoge los datos medidos por sus vecinos y los procesa, obteniendo como resultado un informe de un tamaño más reducido que las medidas iniciales. Otro ejemplo es cuando la red de sensores debe informar de la posición de un vehículo que la atraviesa, utilizando nodos que no se pueden mover de su posición actual.

Estos grupos deben disponer de una infraestructura de clave local, que les permita abrir canales de comunicación seguros entre uno o varios miembros del grupo. Proteger la seguridad de un grupo dentro de una red de sensores que ya se encuentra protegida no es redundante, ya que hay situaciones en las que el grupo necesita de esa protección.

La autenticación del origen es un factor importante dentro de los grupos seguros. Un mensaje dirigido a algunos o todos los miembros del grupo

debe estar debidamente autenticado, o cualquier mensaje que proceda del interior o exterior de la red de sensores puede considerarse, intencionadamente o no, como procedente del grupo. La confidencialidad es también importante, ya que en ciertos escenarios, como la medición de datos dentro de una central nuclear, el grupo puede querer ocultar el intercambio de información y sus resultados finales al resto de la red. Finalmente, la integridad de los mensajes es también esencial, porque sin ella tanto los mensajes de control como las medidas internas del grupo podrían ser atacadas.

Como en la infraestructura de claves vista en el apartado anterior, existen tres factores básicos a resolver a la hora de diseñar la infraestructura de claves de un grupo seguro: almacenamiento, distribución, y mantenimiento de claves. No obstante, proteger a un grupo de nodos es muy distinto a proteger toda la red. Primero, los grupos se crean en la mayoría de los casos de forma dinámica, cuando la estación base lo ordena o cuando ciertas lecturas (ejemplo: un vehículo aproximándose) fuerzan a la red a organizarse a sí misma. En estos casos, las claves del grupo deben ser negociadas y distribuidas automáticamente a todos los (futuros) miembros.

Segundo, los nodos que pertenezcan a un grupo deben ser capaces de guardar todas las claves necesarias para establecer los canales de comunicación seguros, teniendo en cuenta que en casos extremos puede que no haya espacio en memoria para estas claves. Tercero, debido a que los nodos entrarán y saldrán de su grupo local frecuentemente (ejemplo: cuando se está siguiendo un vehículo en el interior de la red de sensores), las operaciones de mantenimiento deben ser seguras para los grupos, en el sentido que un nodo externo no puede entrar en el grupo cuando no esta invitado y un nodo interno no puede abandonar el grupo demasiado pronto. Finalmente, los grupos deben satisfacer dos requerimientos más: "forward security", es decir, que un nodo que abandone el grupo no pueda acceder a las comunicaciones actuales de éste, y túnel seguro (secure tunnel), donde las medidas realizadas por el grupo deben ser leídas única y exclusivamente por la estación base en ciertos escenarios (como por ejemplo plantas nucleares).

Las infraestructuras de clave local no han sido demasiado investigadas en los últimos años, y existen pocas soluciones, la mayoría de ellas costosas en términos de recursos [13]. Una excepción ha sido la protección de grupos creados estáticamente, o clústers, que se configuran antes de la creación de la red, y donde nodos con mayores recursos (denominados "cluster heads") están a cargo de manejar y proteger la seguridad del grupo [14]. Aun así, es necesario desarrollar nuevos esquemas que permitan la creación y mantenimiento de grupos seguros de una forma óptima.

3.4. Routing

Los nodos son capaces de enviar un bit de información, en condiciones óptimas (línea de visión sin obstáculos, máximo gasto de energía, sin interferencias), a un distancia máxima de entre 100 y 300 metros. Esto hace necesario el utilizar algoritmos de encaminamiento, ya que en la mayoría de los casos no es posible enviar un paquete de datos directamente hacia su destino dentro de la red.

El diseño de algoritmos de encaminamiento es una tarea compleja [15]. Es necesario que los paquetes sean capaces de alcanzar cualquiera de los nodos (conectividad) mientras éstos cubren la mayor área posible utilizando sus sensores (cobertura), incluso cuando empiecen a fallar debido a problemas energéticos o de otro calibre (tolerancia a fallos). El algoritmo debería ser también capaz de funcionar con cualquier número de nodos o densidad de la red (escalabilidad) y proveer una calidad de servicio. Al mismo tiempo, los diseñadores deben tratar de reducir al máximo posible los requisitos de memoria, energía y CPU.

La seguridad es otro factor que no puede ignorarse en el diseño de algoritmos de encaminamiento. Cualquier adversario tiene a su disposición una gran variedad de ataques [16] que le permiten manipular a su antojo los caminos de la red, provocando pérdidas, alteraciones o falsificaciones de paquetes. Como ejemplo, es posible redirigir el tráfico de la red hacia un conjunto de nodos anunciándolos como nodos con mejores características, reales o no, de velocidad o conectividad. Es posible también modificar los mensajes de control, o utilizar múltiples identidades en un ataque "sybil".

La infraestructura de claves es útil en la protección de los algoritmos de encaminamiento, ya que permite autenticar a los nodos y proteger la confidencialidad e integridad de los paquetes. Sin embargo, no es suficiente. Tomando el control de un grupo de nodos de la red, un adversario puede modificar cualquier mensaje de control en su propio beneficio. Además, la red puede recibir un ataque de denegación de servicio (DoS) en cualquiera de sus secciones. Es por tanto necesario diseñar algoritmos de encaminamiento que sean robustos ante todos estos tipo de ataques.

Hasta ahora, las investigaciones se han enfocado principalmente en dos áreas: la protección de algoritmos de encaminamiento previamente existentes, tales como la difusión dirigida (directed diffusion [17]), y el descubrimiento de nuevas técnicas para proteger los algoritmos, como por ejemplo los caminos redundantes entre nodos [18] o el descubrimiento y marcado de zonas sin cobertura [19]. Pero la mayoría de los protocolos existentes no tienen en cuenta la seguridad en ninguno de los pasos de su diseño.

Como conclusión, el mayor reto en este área es el de descubrir nuevas técnicas de protección y aplicarlas a nuevos algoritmos, que a la vez que in-

corporan la seguridad como un requisito en todas las fases de su diseño tengan en cuenta los factores esenciales previamente mencionados (conectividad, escalabilidad, etc).

3.5. Agregación de Datos

Dentro de una red de sensores, los nodos generan una inmensa cantidad de datos producto de las mediciones realizadas al entorno. En la mayoría de los casos estos datos deben ser enviados a la estación base, por lo que hay un gran costo, en términos de consumo de energía y ancho de banda, en transportar todos estos datos a través de la red. Sin embargo, ya que los nodos suelen estar densamente distribuidos, los datos procedentes de nodos pertenecientes a una misma zona serán redundantes. El rol de la agregación es el de aprovechar esta situación y resumir todos los datos redundantes en un solo informe, por lo que se reduciría el envío de información hacia la estación base.

Este proceso de agregación es presa fácil de cualquier adversario, incluso aunque la red esté protegida contra ataques hacia la integridad de sus datos. Si un nodo agregador es controlado por un adversario, puede fácilmente ignorar los datos procedentes de sus vecinos y crear un informe falso. Y aún en el caso de que un nodo agregador sea de confianza, éste puede recibir datos manipulados o erróneos.

Utilizando funciones matemáticas que sean resistentes ante ataques internos, es posible defender al nodo agregador ante datos que provengan de nodos manipulados o en mal estado. Utilizando ideas de la teoría estadística, el autor en [20] analizó la robustez de un conjunto de funciones (por ejemplo, demostrando que el mínimo, el máximo, la suma, y la media son funciones inseguras) y propuso algunas herramientas (ej. ignorar valores extremos) para mejorar la robustez de las funciones de agregación.

Existen también soluciones orientadas a descubrir cuando los informes enviados por un nodo agregador están falsificados o no. Una posibilidad consiste en entablar una negociación entre la estación base y el agregador sobre los datos empleados en la creación del informe. Por ejemplo, en [21] la prueba que el agregador debe crear sobre los datos procedentes de sus vecinos se construye sobre un Árbol Hash Merkle.

Existe otra solución que utiliza la densidad de las redes como herramienta, haciendo que los nodos vecinos funcionen como testigos de la agregación. Ellos realizarán los mismos cálculos que el agregador, obteniendo un resultado parecido al estar en la misma zona física. Como ejemplo, en [22] los nodos crean una prueba de sus cálculos utilizando para ello un MAC y una clave secreta compartida con la estación base, de tal forma que el agregador debe enviar a la estación base tanto el informe como las pruebas de los testigos.

Finalmente, es también posible filtrar los paquetes que contienen el informe y las pruebas cuando ambos se encaminan a la estación base, disminuyendo el tráfico generado por informes falsos. En [23], las pruebas creadas por los testigos utilizan una clave procedente de un "key pool", y el agregador las comprime utilizando un filtro de Bloom. En el camino, los nodos que posean una clave del "key pool" pueden comprobar si una prueba está en el interior del filtro de Bloom.

La agregación segura es un campo con muchos interrogantes por resolver. Los protocolos interactivos entre los agregadores y la estación base consumen muchos recursos, y no son escalables sin la presencia de una jerarquía de comprobación de informes. Los sistemas basados en pruebas requieren en la mayoría de los casos de una negociación entre el agregador y los testigos, además de incrementar el tamaño de los informes a enviar. En definitiva, sería necesario disponer de nuevas soluciones que minimicen tanto el número de negociaciones necesarias como el tamaño de los informes, y que introduzcan nuevas técnicas para detectar y eliminar informes falsos más eficazmente.

3.6. Auditoría

Dentro de una red de sensores un usuario solo podrá acceder a la red de adquisición de datos, en la mayoría de los casos, a través de la estación base. Como resultado, cualquier cambio en el estado interno de los nodos (bajo nivel de batería, fallos en el hardware) o de la red pasaría inadvertido. Sería por lo tanto indispensable proporcionar un subsistema de auditoría dentro de la red que permitiera a los usuarios preguntar o recibir informes periódicos acerca de su estado.

Una posible aplicación de ese subsistema de auditoría sería un *Sistema de Detección de Intrusiones* (IDS). Éstos sistemas monitorizan las actividades de la red, recogiendo y analizando datos sobre su comportamiento, con el objetivo de detectar intrusos y alertar al usuario de este hecho. Estos sistemas pueden considerarse, en cierta forma, como una "Segunda Línea de Defensa", que se activa una vez un adversario haya tomado control de ciertas partes de la red.

Un IDS para redes de sensores podría aprovecharse de los conceptos y las técnicas de los IDS desarrollados para redes "Ad Hoc" [24]. Sin embargo, éstas técnicas no pueden aplicarse directamente a las redes de sensores, debido a sus características únicas. Cada nodo de la red no puede realizar de forma completa todas las tareas de detección debido a sus limitaciones de energía y CPU. Además, dado que la densidad de las redes de sensores suele ser alta, sería redundante obligar a todos los nodos a vigilar los paquetes enviados en su vecindario. Por lo tanto, el problema más básico a la hora de desarrollar un IDS es la distribución de las tareas de detección entre los nodos de la

red, problema que actualmente cuenta con algunas soluciones en entornos basados en clusters [25].

Existen también otros problemas aún no resueltos o discutidos en este área. Un IDS debe ser simple y altamente especializado, capaz de analizar y reaccionar ante los problemas que se den en los protocolos de la red. El conjunto de reglas utilizado por los algoritmos de detección debe ser simple y fácil de interpretar, produciendo resultados que consuman poca memoria. Los nodos con tareas de detección deben ser capaces de intercambiar información entre ellos para alcanzar un mejor porcentaje de detección, y las alertas generadas por la arquitectura deben llegar a la estación base lo antes posible.

Cabe mencionar aquí que existen soluciones parciales que son capaces de comprobar la integridad de los nodos de la red, y que podrían ser incorporadas en un IDS. Una de estas soluciones (health monitoring [26]) permite al usuario comprobar si un grupo de nodos se encuentra activo o no. Otro algoritmo, utilizado en [27], analiza fluctuaciones en las mediciones de los nodos utilizando el "Hidden Markov Model" (HMM) para descubrir variaciones inesperadas. Finalmente, es también posible comprobar la integridad del "firmware" de un nodo utilizando técnicas de atestación de código (Code Attestation Techniques [28]).

3.7. Otros Problemas

Una red de sensores necesita de una infraestructura de seguridad que le permita protegerse tanto de los ataques internos como de los ataques externos a la confidencialidad, integridad, y autenticación de los elementos de la red. Sin embargo, esto no es suficiente para resolver determinados problemas que no han sido suficientemente desarrollados en la literatura, tales como la privacidad y la seguridad de agentes móviles.

La privacidad, en determinadas situaciones, es una propiedad esencial. Por ejemplo, en un campo de batalla, sería importante ocultar la localización y las identidades de la estación base y de los nodos que generan información. En contraste, en un escenario de rescate (ej: terremoto), localizar a los nodos (ej: perros) es algo absolutamente necesario.

Existen tres tipos de amenazas contra la privacidad [29]. Si un adversario es capaz de determinar el sentido de un mensaje sólo por su existencia y por el contexto del entorno que rodea a la red, existe una *amenaza contra la privacidad de contenido* (content privacy threat). Si un adversario es capaz de deducir las identidades de los nodos que se están comunicando, existe una *amenaza contra la privacidad de identidad* (identity privacy threat). Y si el adversario es capaz de deducir o aproximar la localización de uno de los nodos que participan en una comunicación existe una *amenaza contra la privacidad de localización* (location privacy threat).

Existen algunos estudios sobre las amenazas de privacidad de localización y contenido [29] que exploran la privacidad de algunos protocolos de encaminamiento. Pero en general, la privacidad en un entorno de redes de sensores es un campo inexplorado, en el que sería importante descubrir e investigar los escenarios en los que existe una amenaza contra la privacidad.

Dado que las redes de sensores están dando sus primeros pasos, existen algunas aplicaciones cuyos requerimientos de seguridad no están aún investigados. Un ejemplo es el área de los agentes móviles [30], que proporciona una interesante herramienta para procesos de computación distribuida. No obstante, cualquier adversario puede ser capaz tanto de incluir en la red un agente malicioso como de modificar los resultados almacenados dentro del agente. Por lo tanto, sería necesario tanto investigar como proveer integridad de código e integridad de resultados dentro de un entorno tan limitado como una red de sensores.

4. Conclusión

La seguridad en redes de sensores es un campo de investigación que está creciendo rápidamente y alcanzando resultados que pueden aplicarse en escenarios la vida real. Durante este crecimiento se ha pasado de un entorno completamente inseguro a disponer de algoritmos básicos, arquitecturas, y herramientas de seguridad.

No obstante, este campo de investigación esta lejos de considerarse maduro. La criptografía de clave pública y los sistemas de detección de intrusiones son técnicas aplicadas recientemente en las redes de sensores. Es también necesario desarrollar algoritmos seguros de encaminamiento que proporcionen conectividad, cobertura y tolerancia a fallos. Finalmente, los algoritmos de agregación de datos deberían ser mas óptimos y seguros, y la privacidad de los flujos de datos debería tomarse en cuenta.

Otras áreas de seguridad en desarrollo en el campo de las redes de sensores [31] son la capacidad de los nodos de aguantar ataques físicos, la optimización de las infraestructuras de seguridad en términos de recursos (energía, tiempo de computación), la detección y reacción ante ataques de denegación de servicio, y la discusión sobre los problemas de privacidad social en las redes de sensores. Finalmente, existen áreas mínimamente desarrolladas que requieren de especial atención, como la protección de redes de sensores con nodos móviles o con múltiples estaciones base, o la medición de la confianza (trust) existente entre nodos.

Referencias

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci. *Wireless sensor networks:*

a survey. Computer Networks, 38(4), Marzo 2002.

[2] Crossbow Technology, Inc. *MICA2 and MICAz, Wireless Measurement Systems*. http://www.xbow.com.

[3] IEEE Standard, 802.15.4-2003. *Wireless medium access control and physical layer specifications for low-rate wireless personal area networks*. Mayo 2003, ISBN 0-7381-3677-5.

[4] C. Karlof, N. Sastry, D. Wagner. *TinySec: a link layer security architecture for wireless sensor networks*. Proceedings of 2nd International Conference on Embedded Networked Sensor Systems (SensSys'04), Noviembre 2004.

[5] N. Sastry, D. Wagner. *Security considerations for IEEE 802.15.4 networks*. Proceedings of 2004 ACM Workshop on Wireless security (Wise'04), Octubre 2004.

[6] D. J. Malan, M. Welsh, M. D. Smith. *A public-key infrastructure for key distribution in TinyOS based on elliptic curve cryptography*. Proceedings of 1st IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (Secon'04), Octubre 2004.

[7] L. Eschenauer, V. D. Gligor. *A key-management scheme for distributed sensor networks*. Proceedings of 9th ACM Conference on Computer and Communications Security (CCS'02), Noviembre 2002.

[8] J. Lee, D. R. Stinson. *Deterministic key predistribution schemes for distributed sensor networks*. Proceedings of 11th Annual Workshop on Selected Areas in Cryptography (SAC'04), Agosto 2004.

[9] B. Yener, S. A. Camtepe. *Combinatorial design of key distribution mechanisms for wireless sensor networks*. Proceedings of 9th European Symposium On Research in Computer Security (ESORICS'04), Septiembre 2004.

[10] W. Du, J. Deng, Y. S. Han, S. Chen, P. K. Varshney. *A key management scheme for wireless sensor networks using deployment knowledge*. Proceedings of IEEE INFOCOM'04, Marzo 2004.

[11] A. Perrig, R. Szewczyk, V. Wen, D. Cullar, J. D. Tygar. *SPINS: Security protocols for sensor networks*. Proceedings of 7th International Conference on Mobile Computing and Networking (MOBICOM'01), Julio 2001.

[12] R. Anderson, H. Chan, A. Perrig. *Key infection: smart trust for smart dust*. Proceedings of 12th IEEE International Conference on Network Protocols (ICNP'04), Octubre 2004.

[13] J. Zachari. *A decentralized approach to secure group membership testing in distributed sensor networks*. Proceedings of 2003 Military Communications Conference (MILCOM 2003), Octubre 2003.

[14] Y. W. Law, R. Corin, S. Etalle, P. H. Hartel. *A formally verified decentralized key management architecture for wireless sensor networks*. Proceedings of 2003 Personal Wireless Communications (PWC'03), IFIP WG 6.8 - Mobile and Wireless Communications. Septiembre 2003.

[15] J. N. Al-Karaki, A. E. Kamal. *Routing techniques in wireless sensor networks: a Survey*. IEEE Wireless Communications, Vol 11(6), pag 6-28, Diciembre 2004.

[16] C. Karlof, D. Wagner. *Secure routing in wireless sensor networks: attacks and countermeasures*. Proceedings of 1st IEEE International Workshop on Sensor Network Protocols and Applications, Mayo 2003.

[17] R. Di Pietro, L. V. Mancini, Y. W. Law, S. Etalle, P. Havinga. *LKHW: A directed diffusion-based secure multicast scheme for wireless sensor networks*. Proceedings of 32nd International Conference on Parallel Processing Workshops (ICPP'03), Octubre 2003.

[18] J. Deng, R. Han, S. Mishra. *A performance evaluation of intrusion-tolerant routing in wireless sensor networks*. Proceedings of 2nd IEEE International Workshop on Information Processing in Sensor Networks (IPSN'03), Abril 2003.

[19] Q. Fang, J. Gao, L. J. Guibas. *Locating and bypassing routing holes in sensor networks*. Proceedings of IEEE INFOCOM'04, Marzo 2004.

[20] D. Wagner. *Resilient aggregation in sensor networks*. Proceedings of 2nd ACM workshop on Security of Ad Hoc and Sensor Networks (SANS'04), Octubre 2004.

[21] B. Przydatek, D. Song, A. Perrig. *SIA: Secure information aggregation in sensor networks*. Proceedings of 1st International Conference on Embedded Networked Sensor Systems (SenSys'03), Noviembre 2003.

[22] W. Du, J. Deng, Y. S. Han, P. K. Varshney. *A witness-based approach for data fusion assurance in wireless sensor networks*. Proceedings of GLOBECOM'03, Diciembre 2003.

[23] F. Ye, H. Luo, S. Lu, L. Zhang. *Statistical en-route filtering of injected false data in sensor networks*. Proceedings of IEEE INFOCOM'04), Marzo 2004.

[24] P. Brutch, C. Ko. *Challenges in intrusion detection for wireless ad hoc networks*. Proceedings of 2003 Symposium on Applications and the Internet Workshops (SAINT'03 Workshops), Enero 2003.

[25] F. Anjum, D. Subhadrabandhu, S. Sarkar, R. Shetty. *On optimal placement of Intrusion Detection Modules in Sensor Networks*. Proceedings of the 1st International Conference on Broadband Networks, Octubre 2004.

[26] C. Hsin, M. Liu. *A Distributed monitoring mechanism for wireless sensor networks*. Proceedings of 2002 ACM Workshop on Wireless Security (WiSe'02), Septiembre 2002.

[27] S. S. Doumit, D. P. Agrawal. *Self-organized criticaly & stochastic learning based intrusion detection system for wireless sensor networks*. Proceedings of 2003 Military Communications Conference (MILCOM'03), Octubre 2003.

[28] A. Seshandri, A. Perrig, L. Van Doorn, P. Khosla. *SWATT: software-based attestation for embedded devices*. Proceedings of 2004 IEEE Symposium on Security and Privacy (S&P'04), Mayo 2004.

[29] C. Ozturk, Y. Zhang, W. Trappe, M. Ott. *Source-location privacy for networks of energy-constrained sensors*. Proceedings of 2nd IEEE Workshop on Software Technologies for Future Embedded and Ubiquitous Systems (WSTFEUS'04), Mayo 2004.

[30] H. Qi, Y. Xu, X. Wang. *Mobile-agent-based collaborative signal and information processing in sensor networks*. Proceedings of the IEEE, 91(8)1172-1183, Agosto 2003.

[31] E. Shi, A. Perrig. *Designing Secure Sensor Networks*. IEEE Wireless Communications, 11(6)38-43, Diciembre 2004.

Construyendo Caminos de Certificación Mediante Cadenas de Hash

Cristina Satizábal[1,2], Rafael Páez[1], Jordi Forné[1]
[1]Departamento de Ingeniería Telemática. Universidad Politécnica de Cataluña
C/Jordi Girona 1-3, C3
08034 - Barcelona (España)
Teléfono: 93 401 59 94 Fax: 93 401 59 81
E-mail: {isabelcs, rpaez, jforne}@entel.upc.es
[2]Facultad de Ingenierías y Arquitectura. Universidad de Pamplona
Km 1 vía a Bucaramanga
Pamplona (Colombia)

Abstract. *Certification path validation is one of the most complex processes of a PKI (Public Key Infrastructure). This demands some storage and processing capacities to the verifier that can exceed the capabilities of some devices, such as mobile phones, smart cards and PDAs. Since, most of the computational cost of this process is derived from the signature verification operations; in this paper we reduce the number of such operations with the use of hash chains to establish the trust relationship among the different entities of a hierarchical PKI. Thus, the integrity of the certificates and their membership of certain certification path are determined by means of simple hash operations.*

1 Introducción

El uso creciente de Internet para realizar todo tipo de operaciones comerciales y de negocios, hace necesaria una infraestructura que garantice el transporte seguro de la información y establezca una relación de confianza entre las entidades participantes. La Infraestructura de Clave Pública (PKI) [1] se encarga de ello, a través de certificados de clave pública (PKCs) que vinculan una clave pública con la identidad de su propietario. Sin embargo, PKI no ha sido ampliamente aceptada debido a su costo, inflexibilidad y difícil manejo.

Uno de los procesos que reviste mayor complejidad en PKI es la validación de caminos de certificación. Este proceso se torna aún más complejo cuando crece la infraestructura, y con ella la longitud de los caminos, lo que supone un mayor trabajo para el verificador y un incremento en sus capacidades de cálculo y almacenamiento. Estos requerimientos no pueden ser soportados algunas veces por dispositivos cuyas capacidades son limitadas como los terminales móviles.

En [2], se ha evaluado el coste computacional de las operaciones criptográficas que realiza un verificador al validar un camino de certificación y aunque este coste es razonable para la PDA considerada [3], llega a ser un aspecto crítico para dispositivos de menor capacidad, como es el caso de las tarjetas inteligentes y los teléfonos móviles, donde este tipo de operaciones requieren mucho más tiempo (ver [4]).

Hasta el momento se han presentado varias propuestas que contribuyen a simplificar la labor del verificador.

Levi y Caglayan [5] propusieron el uso de caminos de certificados anidados (Nested Certificate Paths) para mejorar el desempeño y flexibilidad de los certificados clásicos y simplificar el proceso de validación de caminos. Aunque este método reduce el número de verificaciones criptográficas, su desventaja es el gran número de certificados anidados que deben expedir las NCAs (Nested Certificate Authorities) para construir los caminos en la extensa red de certificados. Esto incrementa la complejidad de la infraestructura y dificulta su gestión.

Por otra parte, Brian Hunter[6] simplifica el uso de PKI, desde el punto de vista del cliente, trasladando las operaciones complejas de los clientes a los servidores. Cada servidor cuenta con una caché local de repositorios y caminos de certificación previos, para que las validaciones posteriores puedan ser significativamente más rápidas. Sin embargo, el camino buscado no siempre esta almacenado en la caché, por lo que el servidor debe realizar todo el proceso de validación de caminos, que es bastante complejo.

En este artículo se utilizan cadenas de hash para establecer la relación de confianza entre las diferentes entidades de una PKI jerárquica, lo que hemos llamado TRUTHC (Trust Relationship Using Two Hash Chains). El uso de cadenas de hash contribuye a disminuir el número de operaciones de verificación de firma durante el proceso de validación de caminos y por tanto el coste computacional del verificador. En la sección 2 se aclara el concepto de camino de certificación y cadena de hash. Además, se describe el proceso de validación de un camino de certificación y las características de la arquitectura jerárquica. La sección 3 especifica el escenario en el cual nos centramos. En la sección 4 se describe

TRUTHC, el método usado para construir caminos de certificación a través de cadenas de hash y sus ventajas. En la sección 5 se compara el coste computacional de los procesos de expedición de certificados y verificación de firma en una PKI típica y el coste obtenido con TRUTHC. Por último, la sección 6 presenta las conclusiones de este estudio.

2 Estado del Arte

2.1 Caminos de Certificación y su Validación

Un camino de certificación [7] es una cadena de certificados de clave pública, que le permite a un usuario determinado obtener la clave pública de otro.

El objetivo principal de una validación de caminos es verificar el vínculo que existe entre el sujeto y la clave pública de un certificado. Para poder confiar en dicha clave pública, el verificador debe chequear la firma y el estado de validez de cada certificado en el camino de certificación de la entidad objetivo.

La base de confianza (trust anchor en inglés) es la clave pública de CA (Autoridad de Certificación) usada por la aplicación del cliente como punto de partida para toda validación de certificados. Por tanto, el camino inicia en la base de confianza del verificador y termina en la clave pública de CA requerida para validar el certificado de la entidad objetivo. La longitud del camino es igual al número de CAs en el camino más uno: un certificado por cada CA y el certificado de la entidad objetivo.

En general, la validación de un camino de certificación involucra los siguientes pasos:

- *Descubrimiento del camino de certificación:* Consiste en establecer un camino confiable entre el verificador y la entidad objetivo a través de las CAs de la infraestructura, basándose en la relación de confianza que existe entre ellas.

- *Recuperación de Certificados:* Consiste en obtener los certificados que hacen parte del camino de certificación de los repositorios donde se encuentran almacenados.

- *Verificación de firmas digitale*s: Consiste en verificar la validez de la firma digital de cada certificado recuperado. Para ello, se debe :

    1. Descifrar la parte firmada del certificado con la clave pública del emisor de dicho certificado.

    2. Calcular un hash sobre el contenido del certificado.

    3. Comparar los resultados obtenidos en 1 y 2. Si coinciden, la firma es válida.

- *Verificación de validez de los certificados:* Se constata si el certificado recuperado ha expirado o esta revocado. La expiración se comprueba con la fecha de caducidad del certificado y el estado de revocación depende del mecanismo de revocación utilizado(ver [1], [8]).

2.2 Arquitectura Jerárquica

Existen diferentes formas de configurar las CAs de una PKI para que sus usuarios puedan descubrir los caminos de certificación: una sola CA, una jerarquía de CAs, una malla de CAs[9]. Como TRUTHC ha sido diseñado para trabajar en una PKI jerárquica solo describimos este tipo de arquitectura.

En una PKI jerárquica, todos los usuarios confían en la misma CA raíz (RCA), es decir, todos los caminos de certificación inician en la clave pública de la RCA. En general, la CA raíz no expide certificados a usuarios, sino a CAs subordinadas. Cada CA subordinada puede emitir certificados a los usuarios o a otro nivel de CAs subordinadas, si las políticas se lo permiten (Fig. 1).

En una PKI jerárquica la relación de confianza es unidireccional, es decir, una CA subordinada no puede expedir certificados a una CA superior en la jerarquía. Gracias a este tipo de relación, los caminos de certificación son fáciles de construir. El camino más largo es igual a la altura del árbol menos uno, ya que el certificado de la base de confianza no se incluye en el camino.

Los problemas de la PKI jerárquica se deben principalmente a la confianza en un solo punto, ya que el compromiso de la clave privada de la CA raíz compromete a toda la PKI. Además, la transición de un conjunto aislado de CAs a una PKI jerárquica puede ser logísticamente impracticable pues todos los usuarios tendrían que ajustar su base de confianza.

2.3 Cadenas de Hash

Una cadena de hash [10] es una lista de valores $y_1, y_2, ..., y_m$ unidos criptográficamente, donde m es la longitud de la cadena. Estas cadenas se obtienen aplicando repetidamente una función de hash H a una semilla secreta x.

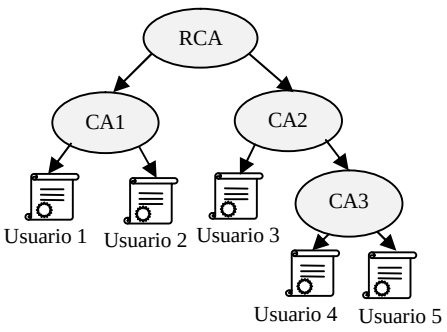


Figura 1: Arquitectura Jerárquica

$$y_1 = H(x)$$
$$y_2 = H(y_1)$$
$$\vdots$$
$$y_m = H(y_{m-1})$$

La función de hash H debe ser una función unidireccional es decir:

1. Dado *x* es fácil calcular H*(x)*

2. Dado un *y* no es posible calcular un *x* tal que *y*=H*(x)*

De esta manera, dado un valor y_i de la cadena no es posible calcular los valores previos.

Además, H puede ser libre de colisión, lo que significa que es computacionalmente imposible encontrar un par (*x, z*) tal que H*(x)*=H*(z)*.

3 Escenario de Estudio

Dos usuarios U y V hacen parte de la misma PKI jerárquica. RCA es la base de confianza de la arquitectura, por tanto, U y V conocen su clave pública (PK_{RCA}) desde el momento en que se registraron en la PKI. Si el usuario V recibe un mensaje M firmado por U y quiere verificar la firma de dicho mensaje, debe primero validar el camino de certificación de U. En Fig. 2 se puede observar con mayor claridad este escenario.

Las flechas indican la relación de confianza que existe entre las diferentes entidades de la PKI. Así, RCA expide los certificados $CERT_{CA1}$ y $CERT_{CA7}$, dirigidos a las autoridades CA1 y CA7 respectivamente; CA1 le expide el certificado a CA2 ($CERT_{CA2}$) y así sucesivamente hasta que la autoridad CA_{L-1} le expide el certificado al usuario U ($CERT_U$), donde *L* es la longitud del camino de certificación. De manera similar se forma el camino de certificación de V.

El círculo con flecha al lado izquierdo de RCA significa que esta autoridad se expide su propio certificado, es decir, $CERT_{RCA}$ es un certificado autoafirmado[1].

En la Tabla 1 se especifica la notación utilizada en este artículo.

El certificado $CERT_X$ se compone de un contenido Cnt_X y la firma sobre ese contenido Sig_X:

$$CERT_X = Cnt_X + Sig_X$$
$$Sig_X = SK_{CAi}(H(Cnt_X))$$

Donde: CAi es la autoridad que expidió dicho certificado.

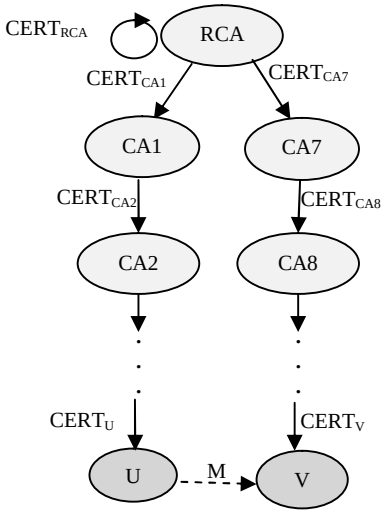


Figura 2: Escenario de Estudio

Tabla 1: Notación Utilizada

Notación	Significado
PK_X	Clave pública de X
SK_X	Clave privada de X
H(D)	Hash sobre la estructura de datos D
$SK_X(D)$	Cifrado/descifrado con la clave privada de la autoridad X sobre la estructura de datos D
$PK_X(D)$	Cifrado/descifrado con la clave pública de la autoridad X sobre la estructura de datos D. Si $D=SK_X(D')$, se espera que retorne D'.
Cnt_X	Contenido del certificado de la entidad X
Sig_X	Firma sobre el contenido Cnt_X
$CERT_X$	Certificado de la entidad X
s_{RCA}	Semilla secreta aleatoria de RCA
n_X	Semilla secreta de la autoridad X
N_X	Semilla encapsulada de la autoridad X.
SN_X	Número serial del certificado $CERT_X$
L	Longitud del camino de certificación.
h_X	Valor de chequeo de integridad asociado a la autoridad X
OP_{hash}	Número de operaciones de hash
T_{hash}	Tiempo de ejecución de una operación de hash
OP_{en}	Número de operaciones de cifrado
T_{en}	Tiempo de ejecución de una operación de cifrado
OP_{dec}	Número de operaciones de descifrado
T_{dec}	Tiempo de ejecución de una operación descifrado
OP_{sig}	Número de operaciones de firma
T_{sig}	Tiempo de ejecución de una operación de firma
OP_{ver}	Número de operaciones de verificación
T_{ver}	Tiempo de ejecución de una operación de verificación
COST	Coste Computacional

Así, si V quiere verificar el camino de certificación del usuario U - $CERT_{CA1}$, $CERT_{CA2}$,…, $CERT_U$ - debe comprobar primero la firma de $CERT_{CA1}$, para

lo cual calcula el hash de Cnt_{CA1} y luego realiza una operación de verificación sobre Sig_{CA1} con la clave pública de la base de confianza (PK_{RCA}). Si los dos resultados coinciden, se puede confiar en el contenido de $CERT_{CA1}$ y por tanto en su clave pública PK_{CA1}. Con esta clave pública se verifica la firma del siguiente certificado en el camino ($CERT_{CA2}$) y así sucesivamente hasta obtener la clave pública del usuario U (PK_U). Si la longitud del camino de certificación es *L*, se requerirán *L* operaciones de hash y *L* operaciones de verificación para comprobar la firma de todos los certificados en el camino. Fig. 3 muestra la forma en que se relacionan los certificados de un camino de certificación.

Cuando se verifica la firma de los certificados que hacen parte de un camino de certificación se persiguen básicamente dos objetivos:

1. Asegurar la integridad de los certificados que hacen parte del camino. Esto se consigue mediante las operaciones de hash

2. Verificar el origen de los certificados o la relación de confianza que existe entre las diferentes entidades que hacen parte del camino. Esto se hace básicamente a través de las operaciones de verificación.

Ya que el coste de las operaciones de verificación es mucho más alto que el de las operaciones de hash, se puede reducir el coste computacional del verificador si se disminuye el número de operaciones de verificación durante el proceso de validación de caminos estableciendo una relación de confianza diferente entre las entidades que hacen parte de una PKI jerárquica.

4 TRUTHC (Trust Relationship Using Two Hash Chains)

TRUTHC establece la relación de confianza entre las entidades de una PKI jerárquica a través de dos cadenas de hash: una encadena las semillas secretas de las CAs y la otra encadena los certificados de cada camino. Esto reemplaza las operaciones de verificación por operaciones de hash en el proceso de validación de caminos de certificación, lo que reduce el coste computacional.

Para describir esta propuesta se utiliza la metodología usada por Karjoth et al, en [11]

4.1 Expedición de Certificados

TRUTHC extiende el proceso de expedición de certificados típico [1]. La relación de encadenamiento de las semillas y de los certificados, el cifrado de las semillas y el protocolo se definen así:

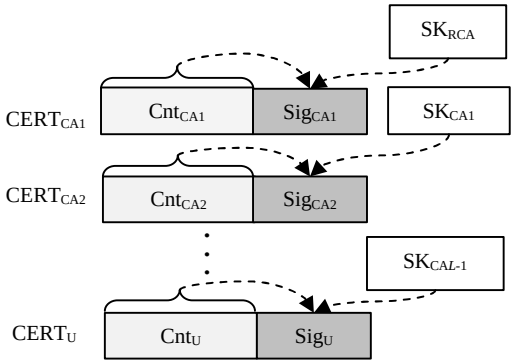


Figura 3: Camino de Certificación de U

Relación de encadenamiento

$$n_{RCA} = H(s_{RCA})$$
$$n_{CA1} = H(n_{RCA}, SN_{CA1})$$
$$n_{CAi} = H(n_{CAi-1}, SN_{CAi}),\ 2 \le i \le L-1$$

$$h_{CA1} = H(n_{RCA}, Cnt_{CA1})$$
$$h_{CAi} = H(h_{CAi-1}, n_{CAi-1}, Cnt_{CAi}),\ 2 \le i \le L-1$$
$$h_U = H(h_{CAL-1}, n_{CAL-1}, Cnt_U)$$

Semilla Cifrada

$$N_{CAi} = PK_{CAi}(n_{CAi}),\ 1 \le i \le L-1$$

Protocolo

$CAi \rightarrow CAi+1$: $CERT_{RCA}$, $CERT_{CAi+1}$, h_{CAi+1}, N_{CAi+1}
$CA_{L-1} \rightarrow U$: $CERT_{RCA}$, $CERT_U$, h_U

Donde: s_{RCA} es la semilla secreta aleatoria de RCA.

Se asume que la función de hash H es libre de colisión.

El protocolo inicia cuando RCA elige la semilla secreta aleatoria s_{RCA} y obtiene n_{RCA}. Luego, RCA expide el certificado $CERT_{CA1}$ y calcula h_{CA1} y n_{CA1} utilizando su semilla secreta n_{RCA}. Más tarde, RCA le envía a CA1: el certificado de la base de confianza $CERT_{RCA}$, el certificado expedido $CERT_{CA1}$, el valor de chequeo de integridad h_{CA1} y la semilla n_{CA1} cifrada con la clave pública de CA1, de manera que solo CA1 pueda descifrarla.

Después, CA1 le expide un certificado a CA2, y realiza las siguientes operaciones:

$$n_{CA1} = SK_{CA1}(PK_{CA1}(n_{CA1}))$$
$$h_{CA2} = H(h_{CA1}, n_{CA1}, Cnt_{CA2})$$
$$n_{CA2} = H(n_{CA1}, SN_{CA2})$$
$$N_{CA2} = PK_{CA2}(n_{CA2})$$

De manera que CA1 le envía a CA2: $CERT_{RCA}$, $CERT_{CA2}$, h_{CA2}, N_{CA2}

Y así sucesivamente, hasta que el usuario U recibe de su autoridad de certificación CA_{L-1} el certificado de la base de confianza $CERT_{RCA}$, su certificado $CERT_U$ y el valor de chequeo de integridad h_U.

Por tanto, se crea una cadena de hash con las semillas secretas (n_{CAi}) y el número serial de los certificados (SN_{CAi}):

$$n_{RCA} = H(s_{RCA})$$
$$n_{CA1} = H(n_{RCA}, SN_{CA1})$$
$$n_{CA2} = H(n_{CA1}, SN_{CA2})$$
$$\vdots$$
$$n_{CAL-1} = H(n_{CAL-2}, SN_{CAL-1})$$

Donde: $L - 1$ es el número de CAs subordinadas en el camino de certificación.

Y otra cadena de hash con los valores de chequeo de integridad (h_{CAi}), las semillas secretas (n_{CAi}) y el contenido de los certificados (Cnt_{CAi}):

$$h_{CA1} = H(n_{RCA}, Cnt_{CA1})$$
$$h_{CA2} = H(h_{CA1}, n_{CA1}, Cnt_{CA2})$$
$$\vdots$$
$$h_{CAL-1} = H(h_{CAL-2}, n_{CAL-2}, Cnt_{CAL-1})$$
$$h_U = H(h_{CAL-1}, n_{CAL-1}, Cnt_U)$$

Fig. 4 muestra la relación que se establece ahora entre los certificados.

4.2 Verificación de Certificados

Adicionamos una tercera parte de confianza (TTP) a la PKI llamada Autoridad de Verificación (VA) (Fig. 5). VA verifica la integridad de los certificados y la relación de confianza entre las entidades que hacen parte del camino de certificación.

RCA expide el certificado de VA ($CERT_{VA}$) y le envía: el certificado de la base de confianza $CERT_{RCA}$, el certificado $CERT_{VA}$ y la semilla n_{RCA} cifrada con la clave pública PK_{VA} para asegurar su confidencialidad.

Semilla Cifrada

$$N_{VA} = PK_{VA}(n_{RCA})$$

Protocolo

$$CA \rightarrow VA: CERT_{RCA}, CERT_{VA}, N_{VA}$$

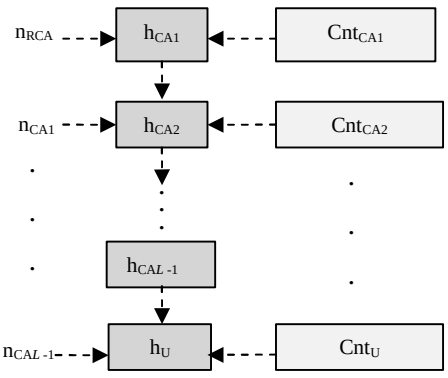


Figura 4: Encadenamiento de Certificados Mediante Cadenas de Hash

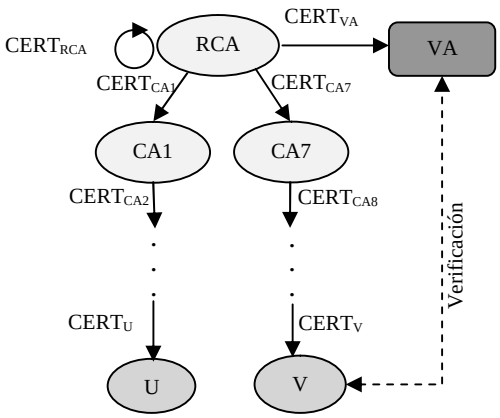


Figura 5: Modelo de Verificación

Para verificar el camino de certificación del usuario U:

1. El usuario V obtiene h_U y el certificado $CERT_U$.
2. V envía a VA el certificado $CERT_U$
3. VA recupera los demás certificados en el camino de certificación de U: $CERT_{CA1}$, $CERT_{CA2}$,..., $CERT_{L-1.}$
4. VA calcula $h'_{U,}$ utilizando las ecuaciones de relación de encadenamiento especificadas en la sección 4.1, la semilla secreta n_{RCA} y los certificados del camino de certificación de U.
5. VA envía el valor de chequeo de integridad h'_U al usuario V en una respuesta firmada.
6. El usuario V verifica la firma de la respuesta de VA. Esto implica verificar la firma de $CERT_{VA}$ con la clave pública PK_{RCA} para obtener PK_{VA} y verificar luego la firma de la respuesta.
7. Si h_U y h'_U son iguales, V habrá comprobado la integridad del certificado $CERT_U$ y su pertenencia al camino de certificación de U.

4.3 Integración con los Certificados X.509

El valor de chequeo de integridad h_{CAi} puede incluirse como una extensión más en los certificados X.509, de manera que TRUTHC es compatible con ellos. Sin embargo, es bueno tener en cuenta que VA debe excluir esta extensión cuando calcule el hash sobre el contenido Cnt_{CAi}, de los certificados que hacen parte del camino, para obtener h'_U.

Adicionar esta extensión a un certificado X.509 implicará un leve incremento en su tamaño, por ejemplo, si se usa SHA-1, el tamaño de los valores de hash será solo 20 bytes.

4.4 Propiedades de Seguridad

Confidencialidad de la semilla: Si el sistema es seguro, sólo la autoridad CAi puede descifrar

$N_{CAi}=PK_{CAi}(n_{CAi})$. Así, cada CA de la PKI conocerá su propia semilla y podrá calcular las semillas de sus CAs subordinadas, pero nunca la semilla de las autoridades superiores en la jerarquía.

Integridad de los certificados: Si un atacante modifica el contenido del certificado $CERT_{CAi}$, que pertenece al camino, y deja intacto h_{CAi}, para que se mantenga la relación de encadenamiento con el contenido modificado Cnt'_{CAi} se debe cumplir:

$$H(h_{CAi-1},n_{CAi-1},Cnt'_{CAi})=H(h_{CAi-1},n_{CAi-1},Cnt_{CAi})$$

Pero esto no cumple con la condición de que la función de hash H sea libre de colisión. Por tanto, no es posible modificar el contenido de un certificado en el camino de certificación sin modificar la relación de encadenamiento que existe entre ellos.

Resistencia a la inserción: Si el atacante quiere incluir un nuevo certificado en el camino y el esquema de cifrado es seguro, aunque disponga de h_{CAi}, no podrá obtener el valor de n_{CAi}, necesario para calcular h_{CAi+1}.

Verificabilidad de la integridad de los certificados: Sólo las CAs pueden calcular los valores de chequeo de integridad h_{CAi} de los certificados que expiden gracias a que son las únicas que conocen las semillas secretas n_{CAi}. Por tanto, si el valor h_U recuperado por un verificador V coincide con el valor h'_U que le retorna VA, este verificador puede confiar en la integridad de los certificados recuperados, y en que todos hacen parte del mismo camino de certificación.

Verificabilidad de la integridad de la cadena de hash: Si el atacante modifica el valor de algún h_{CAi} para propiciar un ataque de negación de servicio (DoS), cuando VA encuentre un h_{CAi} erróneo puede verificar la firma del certificado implicado para asegurar la integridad de dicho valor de la cadena.

5 Evaluación

En esta sección se compara el coste computacional de los procesos de expedición de certificados y verificación de firma de una PKI típica y una PKI con TRUTHC. Los cálculos se basan en el número de operaciones criptográficas necesarias para realizar dichos procesos. Para ello se utiliza SHA-1[12] como función de hash y RSA-1024 [13]como algoritmo de clave pública.

Como tiempos de ejecución de las operaciones criptográficas realizadas por las CAs y la VA se toman los especificados en la Tabla 2, que son de un ordenador con procesador Pentium 4 a 2,1GHz y sistema operativo Windows XP SP1[14]. Y como tiempos de estas operaciones para un verificador con terminal móvil, se toman los mostrados en la Tabla 3. Estos son valores obtenidos de una PDA Compaq iPAQ H3630 con procesador StrongARM a 206 MHz y sistema operativo Windows CE Pocket PC 2002[3].

Tabla 2: Operaciones Criptográficas de las CAs y VA

Algoritmo	Tiempo de Ejecución
SHA-1	14,029 ns/byte
RSA-1024 Cifrado	0,18ms/operación
RSA-1024 Descifrado	4,77ms/operación
RSA-1024 Firma	4,75ms/operación
RSA-1024 Verificación	0,18ms/operación

Tabla 3: Operaciones Criptográficas del Verificador

Algoritmo	Tiempo de Ejecución
SHA-1	0,19 ms/operación
RSA-1024 Firma	78,25 ms/operación
RSA-1024 Verificación	5,01 ms/operación

Se considera que los certificados de los usuarios son como el certificado cliente del ejemplo D.1 en [15] que ocupa 425 bytes, de los cuales 270 bytes corresponden a su contenido y los certificados de las CAs y VA como el certificado de CA del ejemplo D.2 en [15], que ocupa 473 bytes, de los cuales 318 bytes corresponden a su contenido. Además, como se utiliza SHA-1, el tamaño de cada valor de hash va a ser 20 bytes (160bits). Se considera igualmente que el tamaño de la semilla secreta aleatoria s_{RCA} es 20 bytes.

Para calcular el coste computacional se utiliza la ecuación (1).

$$COST = (OP_{hash}*T_{hash}) + (OP_{en}*T_{en}) + (OP_{dec}*T_{dec})+(OP_{sig}*T_{sig})+(OP_{ver}*T_{ver}) \quad (1)$$

5.1 Expedición de Certificados

Caso 1: PKI Típica. Cuando la CA expide un certificado realiza un hash sobre el contenido del certificado y luego una operación de firma sobre ese hash. En este caso, el coste del proceso de expedición de un certificado es el mismo para la RCA y las CAs subordinadas. La Tabla 4 muestra el coste computacional de este proceso para una PKI típica.

Caso 2: PKI con TRUTHC. Aquí, el coste del proceso de expedición de certificados para la RCA es diferente que el de una CA subordinada. Cuando una CA subordinada expide un certificado a otra CA, además de las operaciones de firma del certificado, esta CA realiza:

- Una operación de descifrado :

  $n_{CAi}=SK_{CAi}(PK_{CAi}(n_{CAi}))$

- Dos operaciones de hash

  $h_{CAi+1} = H(h_{CAi}, n_{CAi}, Cnt_{CAi+1})$
  $n_{CAi+1}= H(n_{CAi}, SN_{CAi+1})$

- Una operación de cifrado

  $N_{CAi+1} = PK_{CAi+1}(n_{CAi+1})$

Tabla 4: Coste Computacional del Proceso de Expedición de Certificados

Caso	Emisor	Coste Computacional
PKI Típica	RCA, CA	4,75ms
TRUTHC	CA	9,71ms
TRUTHC	RCA	4,94ms

La Tabla 4, muestra el coste computacional de este proceso para una CA subordinada, cuando se utiliza TRUTHC.

La RCA, en cambio, además de las operaciones de firma del certificado, realiza:

- Tres operaciones de hash

$$n_{RCA} = H(s_{RCA})$$
$$h_{CAi} = H(n_{RCA}, Cnt_{CAi})$$
$$n_{CAi} = H(n_{RCA}, SN_{CAi})$$

- Una operación de cifrado

$$N_{CAi} = PK_{CAi}(n_{CAi})$$

La Tabla 4 muestra el coste computacional de este proceso para la RCA, cuando se utiliza TRUTHC.

5.2 Verificación de Certificados

Caso 1: PKI Típica. Cuando un verificador comprueba la firma de todos los certificados en el camino realiza en total L operaciones de hash y L operaciones de verificación. El coste computacional de este proceso se muestra en (2).

$$\begin{aligned} COST &= (L*0{,}19*10^{-3})+(L*5{,}01*10^{-3}) \\ &= 5{,}20*10^{-3}*L \end{aligned} \quad (2)$$

Caso 2: PKI con TRUTHC. En este caso, el verificador chequea la firma de la respuesta que le envía VA, lo que implica dos operaciones de hash y dos operaciones de verificación (10,40ms).

Por otra parte, VA realiza:

- Una operación de descifrado

$$n_{RCA} = SK_{VA}(PK_{VA}(n_{RCA}))$$

- L – 1 operaciones de hash

$$n_{CAi} = H(n_{CAi-1}, SN_{CAi})$$

- L operaciones de hash

$$h_{CA1} = H(n_{RCA}, Cnt_{CA1})$$
$$h_{CAi} = H(h_{CAi-1}, n_{CAi-1}, Cnt_{CAi})$$
$$h'_U = (h_{CAL-1}, n_{CAL-1}, Cnt_U)$$

- Una operación de hash y una de firma

$$Sig_{VAresp} = SK_{VA}(H(Cnt_{VAresp}))$$

Donde: VAresp es la respuesta firmada de VA. Por motivos de simplicidad, suponemos que esta respuesta solo contiene h'_U.

El coste computacional de VA se muestra en (3).

$$\begin{aligned} COST &= 4{,}77*10^{-3}+14{,}029*10^{-9}*((L-1)*21+338+ \\ &\quad (L-2)*358+310+20) + 4{,}75*10^{-3} \\ &= 5{,}317*10^{-6}*L + 9{,}52*10^{-3} \end{aligned} \quad (3)$$

Fig. 6 compara el coste computacional del verificador cuando lleva a cabo un proceso de verificación de firma en una PKI típica y el coste computacional de VA, en una PKI con TRUTHC, para diferentes valores de L.

6 Conclusiones

El coste computacional del proceso de validación de caminos de certificación en una PKI jerárquica es algunas veces alto, principalmente para dispositivos con baja capacidad de procesamiento como los teléfonos móviles. De las operaciones que realiza el verificador, las criptográficas son las que requieren un mayor tiempo de procesamiento, especialmente las operaciones de verificación.

En este artículo, se propone TRUTHC, un método que establece la relación de confianza entre las entidades de una PKI jerárquica a través de dos cadenas de hash: una que encadena las semillas secretas de las CAs y otra que encadena los certificados de cada camino de certificación. TRUTHC reduce el número de operaciones de verificación durante el proceso de validación de caminos, y por tanto, el coste computacional de este proceso.

TRUTHC permite verificar la integridad de los certificados y constatar que cada uno de ellos pertenece al mismo camino, efectuando únicamente operaciones de hash. Estas operaciones de hash son realizadas por VA y su coste computacional es más bajo que el coste del verificador en una PKI típica como se muestra en Fig. 6. Además, este coste aumenta muy levemente a medida que se incrementa la longitud del camino de certificación. El coste computacional de verificador en TRUTHC es constante y puede reducirse a la mitad si el verificador conoce la clave pública de VA desde el momento en que entra a formar parte de la PKI.

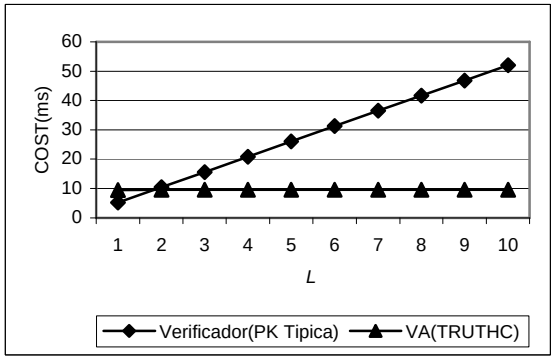


Figura 6: Coste Computacional: Verificador vs. VA

El coste del proceso de expedición de certificados por parte de la CA raíz es similar para una PKI típica y una PKI con TRUTHC, pero es mayor para las CAs subordinadas cuando se utiliza TRUTHC (Tabla 4). Sin embargo, el coste de expedición de un certificado no es muy significativo, gracias a la gran capacidad de procesamiento de las CAs.

Una de las ventajas de este método es su compatibilidad con los certificados X.509, ya que los valores de chequeo de integridad h_{CAi} pueden incorporarse en estos certificados como una extensión.

La seguridad del método propuesto depende en gran medida de la confidencialidad de las semillas n_{CAi}, que sólo deben ser conocidas por las CAs y la VA, por lo que deben ser guardadas de manera segura, como si se tratase de la clave privada de la autoridad. El compromiso de una de estas semillas implica la revocación del certificado de la autoridad propietaria de la semilla y de los certificados expedidos por ella.

Nuestro trabajo futuro se centrará en la definición de la sintaxis de la respuesta firmada de VA y en el mecanismo de actualización de las semillas secretas. También se evaluará la escalabilidad del sistema al incrementar el número de VAs y la eficacia del proceso comunicativo entre las diferentes entidades.

Agradecimientos

Este trabajo ha sido patrocinado por el proyecto ARPA financiado por el Ministerio de Educación y Ciencia (TIC2003-08184-C02-02).

Referencias

[1] ITU-T, "Recommendation X.509: Information Processing Systems - Open Systems Interconnection - The Directory - Authentication Framework (Technical Corrigendum)," 2000.

[2] C. Satizábal and J. Forné, "Revocación de Certificados en la Validación de Caminos de Certificación," in *VIII Reunión Española sobre Criptología y Seguridad de la Información (RECSI 2004)*, vol. 1, *Avances en Criptología y Seguridad de la Información*, A. R. G. Benjamín Ramos Álvarez, Ed. Madrid: Ediciones Díaz de Santos S.A., 2004, pp. 605-613.

[3] P. G. Argyroudis, R. Verma, H. Tewari, and D. O'Mahony, "Performance Analysis of Cryptographic Protocols on Handheld Devices," presented at Third IEEE International Symposium on Network Computing and Applications (NCA'04), 2004.

[4] S. Tillich and J. Grobschädl, "A Survey of Public-Key Cryptography on J2ME-Enabled Mobile Devices," presented at 19th International Symposium on Computer an Information Sciences - ISCIS 2004, Kemer-Antalya, Turkey, 2004.

[5] A. Levi and M. U. Caglayan, "Verification of Classical Certificates Via Nested Certificates and Nested Certificate Paths," in *Eight International Conference on Computer Communications and Networks*, 1999, pp. 242-247.

[6] B. Hunter, "Simplifying PKI Usage Through a Client-Server Architecture and Dynamic Propagation of Certificate Paths and Repository Addresses," presented at 13th International Workshop on Database and Expert System Applications (DEXA'02), 2002.

[7] R. Housley, W. Polk, W. Ford, and D. Solo, "RFC3280 - Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile," April 2002.

[8] M. Myers, R. Ankney, A. Malpani, S. Galperin, and C. Adams, "RFC2560 - X.509 Internet Public Key Infrastructure Online Certificate Status Protocol - OCSP," June 1999.

[9] W. T. Polk and N. E. Hastings, "Bridge Certification Authorities: Connecting B2B Public Key Infrastructures," NIST September 2000.

[10] L. Lamport, "Password Authentication with Insecure Communication," in *Communications of the ACM*, vol. 24, 1981, pp. 770-772.

[11] G. Karjoth, N. Asokan, and C. Gülcü, "Protecting the Computation Results of Free-Roaming Agents," in *Second International Workshop on Mobile Agents (MA'98)*, vol. 1477, *Lecture Notes in Computer Science*. London: Springer-Verlag, 1998, pp. 195-207.

[12] NIST, "Secure Hash Standard." FIPS PUB 180-1, 1995.

[13] R. L. Rivest, A. Shamir, and L. M. Adleman, "A Method for Obtaining Digital Signatures and Public Key Cryptosystems," *Communications of the ACM*, vol. 21(2), pp. 120-126, 1978.

[14] W. Dai, "Crypto ++ 5.2.1 Benchmarks," http://www.eskimo.com/~weidai/benchmarks, 2004.

[15] WAPForum, "WAP Certificate and CRL Profiles," Specification WAP-211-WAPCert-20010522-a, May 2001.

Consensus: Sistema Distribuido de Seguridad para el Testeo Automático de Vulnerabilidades

Guiomar Corral, Agustín Zaballos, Xavier Cadenas, Pete Herzog, Isard Serra
Departamento de Informática. Enginyeria i Arquitectura La Salle. Universitat Ramon Llull
C/ Quatre Camins, 2. 08022 – Barcelona
Teléfono: 93 290 24 23 Fax: 93 211 11 00
E-mail: {guiomar, zaballos, xcadenas, iserra}@salleurl.edu, pete@isecom.org

Abstract*. Nowadays network security has become the main concern when designing, managing and maintaining a network. Any networked system where security and privacy protection of assets is required, needs security experts to protect and control them. There is a need for automated security systems that allow analysts and testers to focus only on critical issues. This paper describes Consensus, a distributed security testing system that analyzes security capabilities, automates vulnerability testing and stores collected data into a database to help security analysis detect vulnerabilities. This scalable vulnerability testing system consists of several interactive probes that provide security operations of vulnerability testing on internal networks, Internet and wireless networks and also expandable to other network technologies due to a modular design. This proposal is based on international best practices for security and follows the methodology and all testing recommendations of the Open Source Security Testing Methodology Manual.*

1 Introducción

Uno de los principales objetivos de la seguridad es proporcionar confidencialidad de la información y disponibilidad de los servicios, intentando a su vez minimizar las vulnerabilidades del sistema donde se encuentra dicha información. Asimismo, el hecho de proteger un determinado recurso no debe afectar a los beneficios que proporcione el acceso a estos a través de la conexión a Internet. A la hora de determinar niveles de seguridad se deben tener en cuenta diferentes factores: conocimiento de la información a proteger, el coste que conlleva el aplicar ciertas medidas de seguridad, garantía de confidencialidad a los usuarios, el mantenimiento del servicio y el acceso a la red sin interrupción.

Un buen análisis de seguridad debe coordinar diferentes fuentes de información para así soportar múltiples modelos de seguridad [1]. Un estudio previo proporcionará la ayuda necesaria para adecuar la seguridad de la arquitectura final. Es imprescindible disponer de una política global de seguridad así como informar a los usuarios de la importancia de ésta. Cabe tener en cuenta que los sensores que controlarán el comportamiento deben estar distribuidos a lo largo de la red a analizar para monitorizar y gestionar el máximo número de comunicaciones. Finalmente se deben planificar testeos de seguridad periódicos con el objetivo de mantener actualizado el nivel de seguridad de la red.

La mejor estrategia a la hora de proteger una red se basa en entender el funcionamiento de los servicios y ser conscientes de las vulnerabilidades subyacentes. Las técnicas de detección de anomalías están ampliamente difundidas en la detección de intrusiones y vulnerabilidades desde su formalización por primera vez en la publicación de Anderson en 1980 [2]. Por esta razón, es muy importante realizar testeos de seguridad de manera periódica. El test de vulnerabilidades es una medida esencial para obtener un marco seguro que cumpla con los requerimientos de la política de seguridad empresarial. De esta forma una organización puede evaluar correctamente el nivel de seguridad de sus redes y, en definitiva, prever cómo un posible atacante puede penetrar en las defensas del sistema y de la red.

Aunque existe información acerca de la realización de testeos sobre una red, no hay una regularización de los procedimientos empleados. Es por ello que, como referencia a la hora de definir dichos procedimientos, se ha seguido, en nuestro caso, el *Open Source Security Testing Methodology Manual* (OSSTMM) [3]. Este manual traza las líneas de conocimiento global sobre las que se debería asentar un análisis de seguridad aplicado a una red [4]. OSSTMM es una metodología abierta que permite verificar de forma fiable la seguridad de una red. Sin embargo, llevar a cabo un test exhaustivo consume una cantidad considerable de recursos humanos y de equipos debido al volumen de información a procesar y al nivel de experiencia del personal requerido para interpretarla. De esta forma la automatización del máximo número de procesos es muy deseable.

Este trabajo describe un nuevo sistema de testeo de vulnerabilidades denominado Consensus que simplifica la ejecución de un testeo OSSTMM mediante la automatización de los procesos asociados. El sistema está formado por varios módulos. Algunos de estos módulos funcionan como sondas que, adecuadamente situadas, llevan a cabo el testeo desde diferentes ubicaciones. El sistema

dispone de una base de datos para almacenar la información referente a los testeos realizados. La información almacenada puede ser revisada por el analista empleando un interfaz *Web*. Los resultados muestran una mejora sustancial en la optimización del tiempo necesario para llevar a cabo los testeos. El presente trabajo se ha llevado a cabo gracias en parte a la subvención PROFIT FIT-360000-2004-81 y a la inestimable colaboración de ISECOM [3].

La organización del artículo es la siguiente: en la sección 2 se presentan los antecedentes sobre seguridad, así como los principales trabajos relacionados con los sistemas de seguridad y sistemas de testeo de vulnerabilidades. En la sección 3 se presentan las especificaciones del Consensus como propuesta para un nuevo sistema automatizado de testeo de vulnerabilidades. En la sección 4 se describe la solución, se detalla la implementación del sistema y se expone el protocolo de comunicaciones. En los apartados 5 y 6 se detallan las experiencias reales, los resultados y, finalmente, las conclusiones.

2 Antecedentes

Mientras que los profesionales de la seguridad en redes se esfuerzan por avanzar y actualizar sus conocimientos sobre las últimas amenazas y vulnerabilidades, los hackers utilizan tecnología de última generación para superar todos los obstáculos. La detección de intrusiones, auditorías y sistemas de registro generan, a menudo, información que no puede ser analizada con efectividad debido a la gran cantidad de datos. La mejor forma de neutralizar dichas vulnerabilidades es desarrollando un conjunto de aplicaciones capaces de revelar puntos débiles de las arquitecturas de las redes corporativas [1].

Actualmente existen múltiples aplicaciones que realizan mediciones de redes, como por ejemplo NIMI y Scriptroute. *National Internet Measurement Infraestructure* (NIMI) es una aplicación capaz de realizar mediciones en Internet. Su arquitectura se basa en un conjunto de servidores y clientes distribuidos en una red y en una máquina de gestión y configuración de mediciones [5]. NIMI está diseñado para efectuar mediciones en un entorno de Internet con la máxima efectividad [6]. Scriptroute es otro sistema de diseño similar y que realiza mediciones como *traceroute*. El sistema es más complejo que NIMI, pero básicamente persigue el mismo objetivo [7]. La filosofía de ambos sistemas es la de proporcionar comunicación y realizar mediciones vía Internet [8]. En cualquier caso el objetivo de dichos sistemas difiere substancialmente de un sistema de testeo de vulnerabilidades, el cual tiene como principal objetivo la automatización de los testeos de seguridad sin que esté específicamente condicionado por el entorno de Internet.

El principal motivo por el que se realiza un test de seguridad a un sistema o a una red es identificar las vulnerabilidades existentes para eliminarlas posteriormente. Debido a que la seguridad requiere una comprobación de los sistemas de forma regular, es importante realizar tests periódicamente y detectar nuevas vulnerabilidades. El CERT Coordination Center describió, en el 1999, solamente 417 vulnerabilidades. Este número se dobló en el año 2000 (1090 vulnerabilidades) y se volvió a doblar en 2001 (2437 vulnerabilidades) para volverse a doblar al año siguiente (4128). En el 2003, se encontraron 3784 vulnerabilidades y el año pasado el número ascendió a 3780 [9].

Los tests de seguridad persiguen, principalmente, tres objetivos. El primero es descubrir defectos de diseño e implementación, así como identificar operaciones que violan las políticas de seguridad. El segundo es el de garantizar que las políticas de seguridad reflejen las necesidades de la empresa. Y el tercero es la evaluación del cumplimiento de la documentación escrita. Un buen test de seguridad completo debería analizar el sistema, las comunicaciones, el acceso físico, el personal, las operaciones y la seguridad administrativa [10]. Aunque todavía no existe ningún estándar que especifique cómo realizar un test de seguridad, OSSTMM está siendo cada vez más aceptado como un estándar de facto. OSSTMM es una metodología diseñada para realizar tests de seguridad y calcular métricas en una red. Se centra en encontrar qué dispositivos requieren ser escaneados, qué hacer antes y después de un test y cómo evaluar los resultados [3]. Su principal objetivo es crear un marco para expertos que utilicen esta metodología y, al mismo tiempo, una guía para realizar testeos exhaustivos a clientes. Un test puede ser considerado un test OSSTMM si es cuantificable, consistente, repetible y exhaustivo [3].

Los proyectos de investigación y los proyectos comerciales desarrollados hasta ahora buscan eliminar vulnerabilidades conocidas y prevenir intrusiones [1,11,12,13]. En estos casos, los profesionales de la seguridad deben instalar y ejecutar diferentes herramientas para realizar un test a una red y encontrar sus vulnerabilidades. Así pues, no sólo puede ser un problema complejo generar y recopilar la información, sino también su almacenamiento y posterior análisis. Por todo ello este trabajo presenta a Consensus, un sistema distribuido que realiza testeos de seguridad de forma automatizada. Este sistema está compuesto por múltiples módulos que ejecutan los testeos [14,15]. Consensus es capaz de realizar testeos de vulnerabilidades sobre distintas tecnologías de redes, como redes corporativas, *De-militarized Zone* (DMZ) y redes inalámbricas entre otras. Además, realiza tests no sólo desde el interior de la propia red corporativa, sino también desde el exterior.

Este sistema proporciona muchas ventajas a las empresas que incentivan la seguridad y la protección de la privacidad. Actualmente, existen herramientas que realizan tests de vulnerabilidades [1,11,12,13],

OSSTMM y la ISO-17799 se están convirtiendo en estándares internacionales y las tecnologías de detección de intrusiones están cada vez más aceptadas [16]. La seguridad y la correlación de datos están saliendo a escena en la investigación con mucha fuerza [17] y se ha intensificado la investigación en la detección "inteligente" de vulnerabilidades. En este contexto es necesario un sistema que integre las nuevas tecnologías con las antiguas y que simplifique la administración y gestión de un Sistema Detector de Vulnerabilidades "inteligente". En este trabajo se ha realizado el primer paso para construir este sistema ideal desarrollando un sistema que automatiza los testeos de vulnerabilidades llamado Consensus.

3 Arquitectura del Consensus

El sistema de test de vulnerabilidades Consensus está compuesto por sistemas interactivos que tienen un diseño base idéntico para proporcionar los mecanismos de test de vulnerabilidades. El sistema automatiza la ejecución de herramientas y minimiza el tiempo necesario para realizar un test que siga la metodología OSSTMM. Actualmente, no existe ningún sistema basado en dicha metodología.

Consensus está basado en productos de seguridad de código abierto, tal como define la propia metodología OSSTMM [3]. Sus características principales son:

- Realización de tests de vulnerabilidad desde una red interna y también desde Internet.
- Automatización de procesos asociados a un test de vulnerabilidad mediante herramientas de libre distribución.
- Configuración y programación de los tests.
- Escalabilidad que permite el uso de múltiples sondas y tecnologías de redes.
- Almacenamiento de resultados en una base de datos para su futuro análisis.
- Visualización de resultados mediante interfaz *Web*.
- Utilización un sistema operativo optimizado para la seguridad.
- Uso de un protocolo de comunicaciones propietario para gestionar las sondas.

3.1 El diseño de Consensus

Un test exhaustivo de seguridad debe considerar todas las perspectivas de una red. Por este motivo una propuesta de automatización de un test tiene que ser modular y, a ser posible, incorporar las nuevas tecnologías, herramientas y actualizaciones a medida que estén disponibles. En este trabajo proponemos Consensus como sistema de test de vulnerabilidades, compuesto por distintos módulos: módulo base, módulo de gestión, módulo de análisis, módulo de base de datos y los módulos sonda. Éstos últimos, adaptados a la tecnología de red empleada: Internet, Intranet, DMZ y *Wireless*. La arquitectura del sistema se muestra en la Fig. 1.

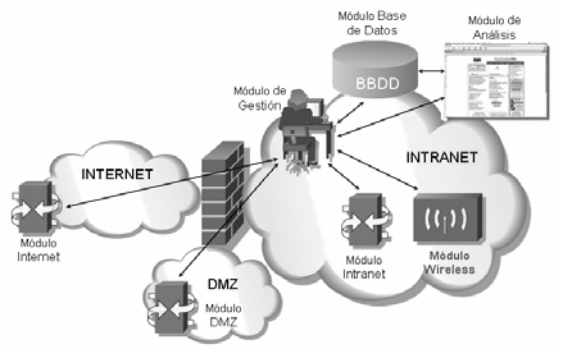


Figura 1: Arquitectura del sistema

El objetivo principal del Módulo Base es ofrecer una plataforma común para configurar testeos, gestionar la seguridad, actualizar el *kernel*, estandarizar el entorno del sistema y asegurar su estabilidad. El resto de módulos funcionan sobre el Módulo Base. El diseño del módulo Base se muestra en la Fig. 2.

El objetivo del Módulo de Gestión es asegurar la estandarización de las actualizaciones y realizar los cambios de configuración desde un entorno centralizado. Establece comunicaciones con todos los subsistemas durante el proceso de test y proporciona el estado de las sondas en tiempo real independientemente de su localización geográfica. Todos los subsistemas del Consensus se controlan desde este módulo mediante un interfaz *Web*.

El diseño del sistema Consensus es escalable ya que permite la configuración de las sondas según el tipo de red o tecnología utilizada. Su arquitectura modular permite la configuración de los tests y de las sondas según sea el tipo de test a realizar (ligero, medio, completo o agresivo). Además, siempre que aparezcan nuevas tecnologías o herramientas de seguridad, el módulo correspondiente puede ser actualizado sin necesidad de reinstalar el sistema existente. Actualmente las sondas disponibles son Internet, Intranet [14], DMZ y *Wireless* [15].

El módulo de Test de Internet, basado en las especificaciones de OSSTMM, se encarga de realizar tests exhaustivos de seguridad a todos los sistemas visibles desde Internet. Este módulo realiza pruebas para descubrir las vulnerabilidades existentes y obtener resultados que informen de las debilidades de la red, del diseño, de los datos y de la estabilidad del sistema. Este módulo es capaz de cifrar los datos privados, sensibles y confidenciales durante todo el proceso del test. Además, es capaz de detectar el uso ilegal o indebido de la red. El módulo DMZ tiene la misma finalidad pero su alcance es, básicamente, los servidores de la empresa. El módulo Intranet es el encargado de realizar el testeo de la red interna, para encontrar las vulnerabilidades que presenta la empresa desde un entorno privilegiado [14]. El módulo de test de *Wireless* permite la realización de tests automatizados a redes inalámbricas así como comprobar el nivel de seguridad de dicha parte de la red, ya que un fallo en la red inalámbrica podría comprometer toda la red corporativa [15].

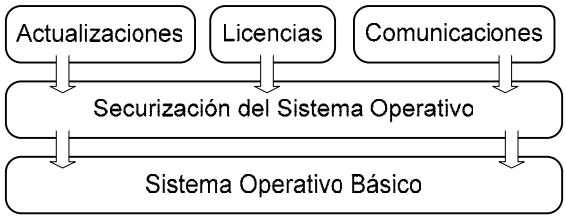


Figura 2: Diagrama del módulo Base

El módulo de Base de Datos almacena toda la información recogida por las sondas. El módulo de Análisis utiliza la base de datos para generar los informes de seguridad. Su finalidad principal es la de recoger la información sobre los sistemas de la red y presentar el informe al analista de seguridad mediante el interfaz *Web*.

4 La implementación del Consensus

Siguiendo los requerimientos del OSSTMM, el sistema Consensus debe utilizar únicamente herramientas de código abierto, no sólo por su bajo coste sino también por su facilidad de adaptación. Por este motivo se ha utilizado una plataforma Linux y la distribución Debian 2.6 como kernel para todos los módulos. Dicho sistema operativo proporciona un entorno estable, configurable y eficiente con un tamaño considerablemente pequeño.

El módulo Base es la parte central del sistema de seguridad; por este motivo es necesario restringir el acceso como administrador con sistemas de cifrado, contraseñas y también un sistema de encriptación basado en claves públicas y privadas. Asimismo, el protocolo de comunicaciones entre módulos utiliza encriptación asimétrica.

El sistema requiere su actualización de forma automática para hacer frente a nuevos riesgos, nuevos módulos, nuevas herramientas y también para mejorar el funcionamiento global del sistema. De no ser así, podría quedar obsoleto en poco tiempo. Para realizar dicha función se ha utilizado la herramienta APT (*Advanced Packaging Tool*).

4.1 Módulos de test del Consensus

Cabe recordar que todos los módulos de test de Consensus (Internet, Intranet, DMZ y *Wireless*) funcionan sobre el módulo Base. El diagrama del funcionamiento de las sondas se muestra en la Fig. 3.

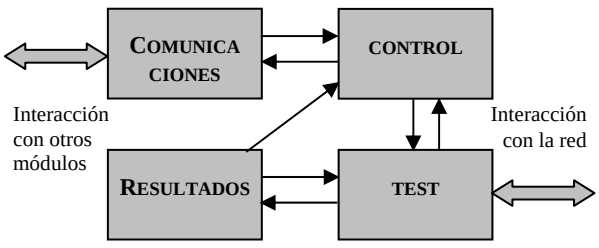


Figura 3: Diagrama de la sonda

Cualquier módulo de Test se comunica únicamente con el módulo de Gestión, el cual envía un fichero de configuración para notificar a la sonda los parámetros del test. El sistema de comunicaciones es propietario y ha sido implementado para controlar las sondas, enviar y recibir los ficheros de resultados. Se describe en la sección 4.3. Cuando el módulo de Test ha finalizado, envía los resultados obtenidos al módulo de Gestión, el cual los introduce en la base de datos.

El bloque 'Control' gestiona los parámetros de configuración y supervisa el test. Además notifica al módulo de Gestión cualquier error producido antes, durante o al finalizar un test. El bloque 'Test' utiliza la información del bloque 'Control' y ejecuta las herramientas de test con los parámetros indicados por el fichero de configuración. Los tests se realizan siguiendo la metodología OSSTMM [3] y acotan el alcance de la red, identifican los servicios que proporciona el sistema, detectan vulnerabilidades conocidas, testean el encaminamiento, ponen a prueba a los *firewalls* e IDSs, emplean mecanismos para descubrir contraseñas e implementan ataques básicos de denegación de servicio (DoS). La Tabla 1 muestra una relación entre las distintas fases del test y las herramientas utilizadas en cada una de ellas.

Tabla 1: Herramientas

DATOS	HERRAMIENTA	FASE OSSTMM
Nombre de host	Nessus	System Service Identification
Sistema operativo	Nmap, Xprobe2	System Service Identification
Estado de los puertos	Nmap	System Service Identification
Servicios de los puertos abiertos	Nmap, THC-Amap	System Service Identification
Vulnerabilidades genéricas	Nessus	Vulnerability Research
Vulnerabilidades específicas: servidores, routers, firewalls…	Nessus, Nikto, md-webscan	Application Testing
Encaminamiento	Irpas	Router Testing
Reglas de Filtrado	Ftester	Access Control Testing
Respuesta a código malicioso	Email, Netcat	Containment Measures Testing
Respuesta IDS	Nmap, Nessus, Nikto	IDS Testing
Contraseñas débiles	John the Ripper	Password Cracking
Respuesta a DoS	Unicornscan, Juno	Denial of Service

4.2 Módulos de Gestión, Base de Datos y Análisis

El módulo de la Base de Datos es el responsable de almacenar los resultados de los testeos. Para su implementación se ha utilizado PostgreSQL, siguiendo los requerimientos de código abierto expuestos en el OSSTMM. Las sondas nunca interactúan directamente con la base de datos; es el módulo de Gestión el encargado de recoger la información de las sondas e introducirla en la base de datos. El módulo de Análisis utiliza dichos datos para generar los informes de seguridad.

El modulo de Gestión es una plataforma centralizada que controla todo el sistema. Interactúa con el profesional de seguridad, emite instrucciones precisas para las sondas y monitoriza los subsistemas y las comunicaciones en todo momento. Además, informa del estado de las sondas en tiempo real cualquiera que sea su localización. Para que el analista de seguridad pueda controlar el sistema, éste dispone de un interfaz *Web*. El acceso a dicho interfaz está protegido mediante contraseña; una vez autenticado, el analista dispondrá de una visión en tiempo real del estado del sistema (Fig. 4).

La configuración de un test se realiza mediante el interfaz *Web* del módulo de Gestión. Para hacerlo es necesario rellenar unos formularios indicando el tipo de sonda, la dirección IP y el dominio, entre otros. Una vez finalizado este proceso, los parámetros de configuración se envían desde el módulo de Gestión hacia las sondas correspondientes (Fig. 5).

Los usuarios del Consensus, las sondas y los tests se pueden configurar desde el módulo de Gestión. El administrador puede añadir, modificar y borrar usuarios; puede añadir y borrar sondas; también puede configurar nuevos tests y borrar los tests que aún no han sido ejecutados, etc. Como se ha especificado anteriormente, para asegurar la confidencialidad de las comunicaciones en este sistema de seguridad, las comunicaciones entre el módulo de Gestión y el resto de módulos emplean mecanismos de cifrado asimétrico.

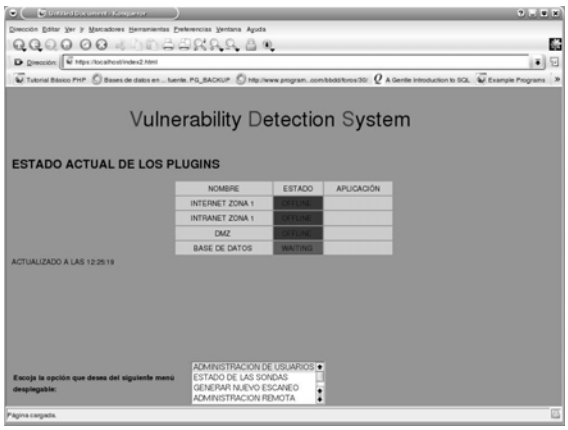


Figura 4: Consola de Gestión – estado de un test

Figura 5: Consola de Gestión – configuración de un test

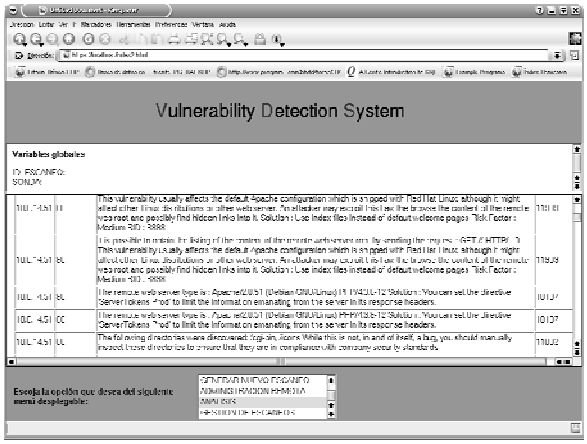


Figura 6: Ejemplo de informe

El módulo de Análisis es el responsable de mostrar toda la información recogida por las sondas al experto en seguridad. Dicha información se encuentra en la base de datos y el usuario puede elegir qué información se debe incluir en el informe final. Se muestra un ejemplo en la Fig. 6.

4.3 El protocolo de comunicaciones

El protocolo de comunicaciones ha sido diseñado especialmente para el sistema Consensus y está basado en el mecanismo de *3-way handshake* [18]. Después de analizar el contenido de la IETF '*Intrusion Detection Exchange Format* (IDWG)' [19], el protocolo de comunicaciones de Consensus sólo incluye las ventajas del IDWG. Las problemáticas encontradas se basan en que IDWG define mensajes con estructuras de tamaño fijo y nuestro sistema necesita enviar información adicional. Además, IDWG envía un mensaje por nodo escaneado y esto podría suponer un problema en una red con muchos ordenadores trasmitiendo al mismo tiempo.

El módulo de Gestión se comunica con los otros módulos utilizando distintos mensajes. El 'State request', por ejemplo, es un mensaje que envía el módulo de Gestión para verificar si las sondas están deshabilitadas, esperan las instrucciones, están

realizando un test o para indicar cuando el testeo ha finalizado. También existe un mensaje especial para realizar un paro de emergencia de una sonda determinada.

El protocolo de comunicaciones no sólo envía y recibe mensajes, también realiza la transferencia de ficheros. Este proceso se ejecuta una vez para enviar el fichero de configuración del test a las sondas y luego para recibir el fichero con los resultados de los tests. El módulo de Gestión recibe un mensaje de la sonda indicando que el test ha finalizado. El protocolo utilizado para este proceso es el SCP (*Secure CoPy*). Este protocolo permite intercambiar ficheros de un modo seguro (*Secure Shell File Transfer*) y es de código abierto. La Fig. 7 muestra el proceso de intercambio de mensajes.

5 Experiencias

En esta sección se exponen las experiencias de la implementación del sistema Consensus en entornos reales. Comenzaremos por una visión global de su estado de implementación actual, seguiremos con una descripción de los escenarios dónde se ha probado el sistema y, por último, mostraremos los resultados de los testeos de vulnerabilidades realizados en entornos controlados.

5.1 Estado de Consensus

El sistema de test de vulnerabilidades Consensus está en fase de implementación beta. El sistema ha sido implementado y probado en entornos reales y es capaz de realizar test automatizados de seguridad que satisfacen las recomendaciones de OSSTMM. Además, recoge los resultados de las pruebas y los almacena en una base de datos centralizada e integrada en el sistema.

Actualmente, el sistema soporta múltiples sondas y arquitecturas de redes; además soporta la realización de tests a distintas máquinas al mismo tiempo. Los test se pueden configurar a través del interfaz *Web* del sistema, el cuál permite al analista de seguridad revisar los resultados una vez finalizado el test de vulnerabilidades.

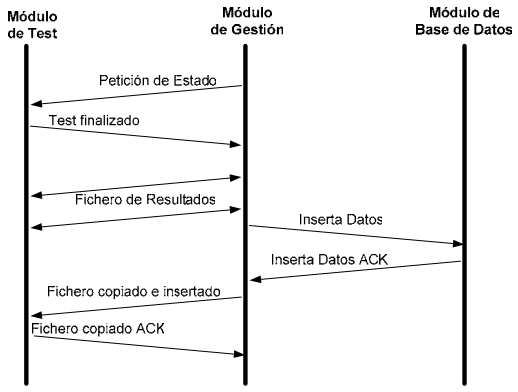


Figura 7: Intercambio de mensajes

5.2 Escenarios de pruebas

El sistema Consensus ha sido probado sobre dos entornos reales. El primer escenario utilizado ha sido un entorno de pruebas de la propia red de la Universidad, debido a su accesibilidad y también a la posibilidad de que el test de seguridad causara interrupciones en dicha red. Es importante mencionar que para realizar un test de vulnerabilidades es necesario un mínimo ancho de banda y los resultados preliminares eran impredecibles en las primeras fases. El segundo entorno de trabajo ha sido una empresa de seguridad. Este escenario ha ayudado a ajustar el tiempo y el ancho de banda necesarios para realizar un test. Los resultados de los testeos son distintos cada vez puesto que el entorno de red es cambiante y, por este motivo, es siempre difícil predecir su duración. En la siguiente sección se muestra una comparativa entre un test realizado manualmente y un test realizado utilizando el sistema Consensus.

5.3 Resultados

Para evaluar el funcionamiento de Consensus se ha realizado un test de vulnerabilidades a una empresa tanto de forma manual como de forma automatizada. Para proteger la privacidad de la empresa, las direcciones IP y los nombres han sido eliminados. Además hay que señalar que el cliente no quiso incluir un análisis de denegación de servicio.

La principal ventaja del sistema automatizado Consensus es que, si bien, en un test manual sólo se contabilizan las horas laborables (horas/hombre), un test automatizado puede ejecutarse en horas y días no laborables. En la Tabla 2 se puede ver una comparación de la duración de un test utilizando los dos métodos. La diferencia entre ambos es la unidad de medida: el test manual está indicado en horas hombre, y el test automatizado, en horas naturales. Este hecho no sólo añade flexibilidad a la ejecución de los pruebas, sino que también reduce considerablemente su coste. Además, la realización de los testeos en horas no laborables puede ayudar a minimizar los efectos sobre la red.

Otro aspecto a tener en cuenta es el tiempo utilizado en la verificación de los testeos. Es común que un test manual requiera tiempo para verificar el correcto funcionamiento de las herramientas de test, analizar los resultados y escribir el informe final. El sistema Consensus, en cambio, ahorra tiempo en realizar el informe ya que los resultados se pueden ver inmediatamente a través del interfaz *Web*. Asimismo, también ahorra tiempo en la instalación y en la verificación de herramientas, puesto que éstas ya están integradas en el sistema global. Es importante destacar que ambos métodos utilizan las mismas herramientas para realizar un test y, por tanto, las pruebas son totalmente equivalentes.

Tabla 2: Duración de un test

Hosts	Manualmente	Consensus
10.0.0.1	2 días hombre	16 horas
10.0.0.2	3 días hombre	22 horas

6 Conclusiones y líneas de futuro

Este artículo ha presentado una nueva propuesta para realizar testeos de vulnerabilidades de forma automatizada y siguiendo el *Open Source Security Testing Methodology Manual* (OSSTMM). Los testeos de vulnerabilidades se están convirtiendo en tareas imprescindibles para mantener una red segura y la experiencia demuestra que una vez conocidas las debilidades de una red es mucho más fácil proteger los sistemas contra *hackers* u otros usuarios malintencionados. Aunque existen programas que realizan tests de vulnerabilidades, no existía hasta ahora ningún sistema capaz de automatizar todos los procesos necesarios para detectar vulnerabilidades en una red interna, en una DMZ, y en una red inalámbrica al mismo tiempo, según especifica el OSSTMM. La modularidad del sistema Consensus ayuda a integrar nuevas tecnologías de redes en cuanto aparezcan y a introducir las mejoras que sean necesarias en cada momento. En este artículo se ha descrito también la arquitectura del sistema y su protocolo de comunicaciones.

Las líneas futuras de esta investigación hacen referencia, principalmente, al módulo de Análisis del sistema. La supervisión de los datos puede mejorarse introduciendo métodos de aprendizaje supervisados y no supervisados [16]. Además, las técnicas de Inteligencia Artificial (IA) pueden ayudar a aprender de las experiencias pasadas, a descubrir patrones de comportamiento, a detectar nuevas vulnerabilidades y a extraer nuevas conclusiones sobre los datos obtenidos en un test [1,20]. La IA ya se ha aplicado en otros campos de la telemática con muy buenos resultados [21].

Agradecimientos

Este proyecto ha sido parcialmente financiado por el proyecto PROFIT FIT-360000-2004-81 del Ministerio de Industria, Turismo y Comercio. También queremos agradecer a Enginyeria i Arquitectura La Salle (EALS), Universitat Ramon Llull por su apoyo a nuestro grupo de investigación de Sistemas Inteligentes (2002 SGR-00/55). Y finalmente, queremos agradecer a Pete Herzog e ISECOM por su apoyo en el proyecto y por el OSSTMM.

Referencias

[1] J. Dawkins, J. Dale. "A Systematic approach to Multi-Stage Network Attack Analysis". Proceedings on the 2nd. IEEE IWIA'04, 0-7695-2117-7/04, 2004.

[2] J.P. Anderson. "Computer Security Threat Monitoring and Surveillance", Technical Report, James P. Anderson Co., Fort Washington, Pennsylvania, April 1980.

[3] P. Herzog. OSSTMM. http://www.isecom.org.

[4] A. Zaballos, G. Corral, I. Serra, J. Abella. "Testing Network Security Using OPNET". OPNETWORK 2003, Washington D.C., August 2003.

[5] V. Paxon, A. Adams, M. Mathis. "Experience with NIMI", PAM 2000 Proceedings, University of Waikato, Hamilton, New Zealand, April 3 and 4, 2000.

[6] V. Paxon, J. Mahdavi, A. Adams, M. Mathis. "An Architecture for Large-Scale Internet Measurement", IEEE Communications, 36 8, 48-54, 1998.

[7] N. Spring, D. Wetherall, T. Andreson. "Scriptroute: A Public Internet Measurement Facility", USENIX Symposium on Internet Technologies and Systems (USITS), 2003.

[8] V. Paxon, J. Mahdavi, A. Adams, M. Mathis. "Creating a Scalable Architecture for Internet Measurement", Inet98 Proceedings, 1998.

[9] CERT Coordination Center Statistics 1998-2004. www.cert.org/stats/cert_stats.html.

[10] J. Wack, M. Tracy, M. Souppaya. "Guideline on Network Security Testing, Recommendations of the NIST"; US Department of Commerce (Computer Security Division),October 2003.

[11] R. Deraison. Nessus: http://www.nessus.org.

[12] Internet Security Systems: http://www.iss.net.

[13] Insecure, Nmap, http://www.insecure.org/nmap.

[14] G. Corral, A. Zaballos, X. Cadenas, A. Grane. "A Distributed Vulnerability Detection System for an Intranet". 39th IEEE International Carnahan Conference on Security Technology. Las Palmas de G. Canaria, Octubre 2005.

[15] G. Corral, X. Cadenas, A. Zaballos, M.Cadenas. "A Distributed Vulnerability Detection System for WLANs". 1st IEEE International Conference on Wireless Internet. Budapest, Julio 2005.

[16] J. Haines, L. Rossey, R. Lippmann, R. Cunningham. "Extending the DARPA Off-Line Intrusion Detection Evaluations", Proceedings of DARPA Information Survivability Conference & Exposition II, Volume: 1, 2001, pp. 35-45.

[17] P. Ning, Y. Cui, D. Reeves. “Analyzing Intensive Intrusion Alerts via Correlation”, Proceedings of the 5th International Symposium on Recent Advances in Intrusion Detection (RAID 2002), LNCS 2516, Zurich, Switzerland, October 2002, pp. 74-94.

[18] W. Richard Stevens, Gary R. Wright. TCP/IP Illustrated Vol.2. Addison-Wesley. November 1994.

[19] IETF. “Intrusion Detection Exchange Format (IDWG)”. 59th IETF Meeting in Seoul, October 2003.

[20] M.Conner, C. Patel, M. Little. "Genetic Algorithm/Artificial Life Evolution of Security Vulnerability Agents". Army Research Laboratory Federal Laboratory 3rd Annual Symposium on Advanced Telecommunications & Information Distribution Research Program. February 1999.

[21] G. Corral, A. Zaballos, J. Camps, J. Garrell. “Prediction and control of short-term congestion in ATM networks using Artificial Intelligence techniques”. Proceedings on the 1st. International Conference on Networking (ICN 2001), LNCS 2094 Springer, Colmar , France, July 2001, pp.648-657.

Aplicabilidad de Redes Neuronales a los Sistemas de Detección de Intrusos

Iván Pau de la Cruz, Esther Gago García, Diego Escarda Tejada, Borja Jiménez Salmerón
Departamento de Ingeniería y Arquitecturas Telemáticas (DIATEL)
EUIT de Telecomunicación. Ctra. Valencia. Km. 7. Campus Sur.
28031 – Madrid
Teléfono: 913365519 Fax: 913367817
E-mail: {ipau,egago,descarda,bjimenez}@diatel.upm.es

Abstract*. IDS (Intrusion Detection Systems) are definitely, in the world of security, one of the most interesting technologies for local area networks protection. However, these systems have limitations in discovering intrusions when such intrusions follow small modifications of the patterns included in the system. Detection effectiveness might be significatively improved through the utilization of tools capable of learning and generalizing from the gathered patterns. Neural networks technologies have demonstrated to be a valid approach in this context. In this article, a prototype is proposed and validated by combining behaviour analyzer with patterns analyzer techniques.*

1 Introducción

La seguridad es un aspecto de vital importancia en los sistemas telemáticos actuales. El valor crítico de gran parte de la información procesada y transmitida en estos sistemas requiere procedimientos fiables que eviten específicamente su robo, pérdida, falsificación o corrupción [1].

Para la protección de esta información y en consecuencia de las aplicaciones que la procesan, existen varios mecanismos posibles. Estos mecanismos pueden estar implementados en la propia aplicación a proteger (de forma independiente o apoyándose en otras entidades) o pueden ser externos a la aplicación. En estos últimos están englobados los sistemas de detección de intrusos (en adelante *IDS*, *Intrusion Detection Systems*).

La seguridad informática, a diferencia de otras disciplinas telemáticas, enfrenta a entidades de estados finitos (procesos, entidades software, etc) con entidades de estados infinitos (inteligencia humana de los atacantes). Debido a esto, los mecanismos orientados a minimizar amenazas están en clara desventaja frente a los generadores de éstas, ya que deben esperarse cualquier estado de los infinitos posibles. La validez de utilizar técnicas propias de la Inteligencia Artificial en sistemas que proveen seguridad ya ha sido constatada por algunos autores [2]. En esta línea, la incorporación de una arquitectura abierta que contemple analizadores de comportamiento y analizadores selectivos de protocolo puede reducir de manera muy efectiva la vulnerabilidad del sistema ante ataques desconocidos.

El principal objetivo del trabajo expuesto es aplicar la tecnología de redes neuronales, que permite realizar tareas de clasificación de patrones con la capacidad de generalización y aprendizaje que no tienen otras tecnologías, en la fase de análisis de los IDS. De esta forma se podrán detectar anomalías que sean variaciones de las conocidas previamente [3].

1.1 Sistemas de Detección de Intrusos

Un IDS es un sistema que realiza una monitorización e interpretación de los eventos ocurridos en determinadas entidades bajo observación. Su principal objetivo es identificar procedimientos no lícitos que puedan comprometer o poner en peligro la integridad, confidencialidad o disponibilidad de un sistema.

Suelen ser elementos pasivos que informan a los administradores del sistema de sus descubrimientos, aunque también pueden tener papeles activos adoptando medidas ante una intrusión. Cuando un IDS es activo se le suele llamar IPS (*Intrusion Prevention System*).

Los IDS constituyen un apoyo muy importante en la gestión de seguridad de cualquier red ya que ayudan a identificar riesgos, amenazas y vulnerabilidades actuales de los sistemas, permitiendo a los administradores adoptar las medidas necesarias en sus políticas de seguridad .

Para llevar a cabo sus funciones, los IDS están compuestos de los siguientes bloques funcionales:

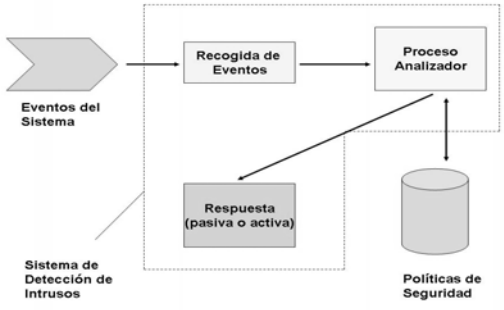


Fig. 1. Bloques funcionales de un IDS

Los eventos son capturados por un sistema de recogida de eventos que se encargará de agruparlos, codificarlos, filtrarlos, etc. Este bloque depende de las posibilidades del sistema en estudio y del tipo de información que se quiere tener en cuenta (a criterio del diseño del IDS). Una vez recogidos los eventos, los formatea y se los pasa al bloque que realiza el proceso analizador.

El analizador estudia los eventos producidos para detectar si hay intentos de intrusión en la red. Para realizar este estudio se pueden usar varias técnicas, aunque comúnmente se utilizan tres: reconocimiento de patrones, captación de frecuencias de acceso y técnicas estadísticas como las usadas en minería de datos [4].

El uso de una o varias de estas técnicas dependerá del objetivo del motor de análisis, es decir, del tipo de detección que se quiere llevar a cabo. Existen dos tipos de detección en los IDS [5]:

- *Usos indebidos*. En este tipo de detección se buscan intentos de intrusión ya conocidos por el sistema. Para implementarlo se suelen usar patrones de intrusión codificados a través de reglas y se busca el emparejamiento de los eventos del sistema con estos patrones.

- *Anomalías*. Por el que se detectan usos poco normales de los recursos del sistema. No vale con reconocimiento de patrones, ya que no hay conocimiento previo de la anomalía. Actualmente se usan técnicas estadísticas.

Finalmente, los resultados del proceso de análisis se envían a un bloque que gestiona la respuesta del sistema (activa o pasiva).

1.2 Talón de Aquiles de los IDS

Las técnicas para la detección de usos indebidos en los IDS están bien definidas y dan resultados aceptables. Sin embargo, detectar únicamente usos indebidos (de los que se tiene conocimiento previo) limita la aplicabilidad de los IDS.

Cuando existen variaciones en los patrones de intrusión, situación muy frecuente, el IDS no es capaz de reconocerlo. Para que el IDS realice su función de forma eficiente, se debe actualizar constantemente sus reglas para facilitar el reconocimiento de variaciones en los patrones. De esta forma, la capacidad de que el sistema responda ante una intrusión dependerá de la frecuencia de actualización de estas reglas. En la mayoría de las organizaciones esta frecuencia no es la adecuada, quedando el sistema expuesto a intrusiones no detectables durante determinados periodos de tiempo (fácilmente predecibles).

Esta situación podría mejorar si el IDS fuera capaz de reconocer anomalías en el sistema. Como se expuso anteriormente, podrían usarse métodos estadísticos para conseguir el reconocimiento de variaciones en patrones de intrusión. El problema de estos métodos es la obtención de patrones y muestras con distribuciones lineales normalizadas y completas. En muchos casos no puede darse esta situación, por lo que se necesita una solución más versátil.

La solución propuesta en el presente artículo es el uso de redes neuronales como medio de apoyo para la detección de anomalías.

1.3 Redes Neuronales

Las redes neuronales son un sistema de procesado de información inspirado en el elemento más básico del sistema nervioso humano: la neurona. Están compuestas de múltiples unidades de proceso independientes conectadas entre sí que se transmiten información a través de conexiones ponderadas. Esa información procesada da lugar a una salida a partir de una determinada entrada [6].

El punto de partida de la utilización de las redes neuronales en la clasificación de patrones de intrusión es el común con otros problemas de clasificación: existe una población que se puede dividir en un cierto número de grupos cuando es observada a la luz de un comportamiento determinado. De este comportamiento es precisamente del que se deduce la asignación. Pero lo realmente interesante es tomar una decisión sobre la pertenencia a un grupo u otro *antes* de que se dé ese comportamiento. En el caso de los IDS el interés radica en decidir *a priori* si un determinado *mensaje o comportamiento* es ataque o no antes de que dañe el sistema.

Toda red neuronal tiene un proceso de diseño de la propia red, en el que se especifica su topología, morfología y los parámetros de funcionamiento, y un proceso de aprendizaje en el que se ajustan los pesos de las conexiones entre unidades para que la red tenga conocimiento del problema al que atiende. A continuación se expondrán estos dos procesos.

Diseño de la red neuronal.

En primer lugar es necesario establecer el número de capas de la red, el número de unidades de procesamiento en cada capa y cómo se conectan unas unidades con otras (arquitectura de la red).

En las redes, la información fluye de una capa a otra de dos formas posibles: en cascada (redes feed-forward) o de forma recurrente (redes feed-backward).

Independientemente de cómo fluya la información en la red, cada unidad recibirá una entrada en función de la cual tendrá una activación y una salida.

Para la función de entrada a la red se suele usar la llamada función lineal (LBF) que es una suma

ponderada de los pesos de las entradas a la unidad. La salida de la red puede ser igual a la entrada, o bien utilizar la función escalón o la sigmoidal, de bastante plausibilidad biológica, que da una salida comprendida entre [0,1]. Hay una variante de la sigmoidal denominada tangente hiperbólica de da salidas comprendidas entre [-1, 1] que permite aplicar polaridades [7].

Aprendizaje de la red.

Una vez diseñada es el momento de realizar el proceso de aprendizaje de la red. El aprendizaje consiste en la presentación de patrones a la red y la subsiguiente modificación de los pesos de las conexiones siguiendo reglas de aprendizaje que traten de optimizar su respuesta minimizando el error.

Existen dos tipos de aprendizaje: supervisado, en el que se presentan los patrones de entrada junto a los patrones de salida deseados y el no supervisado en el que solo se presentan los patrones de entrada, sin los de salida, y se le permite a la red que los organice con alguna regla de auto-organización.

Independientemente del tipo de aprendizaje usado, una parte esencial es la regla de aprendizaje que se utiliza para indicar como se modifican los pesos de las conexiones en función de los datos usados en la entrada. Es necesario tener un número adecuado de patrones para conseguir un buen entrenamiento de la red y forzar a una convergencia en la función de error.

2 Modelo Propuesto

En esta sección se presenta el modelo que se ha definido para llevar a cabo la integración de los IDS y las redes neuronales.

En primer lugar se definirá la arquitectura física del sistema, en la que se podrá ver la topología de la red de comunicación y la elección de los tipos de IDS y de redes neuronales usados. Una vez expuesto este primer nivel arquitectural, se presentará la arquitectura lógica del sistema en la que se comentan los bloques del proceso de análisis del IDS, que será el que incorpore las redes neuronales en su interior.

2.1 Arquitectura física

2.1.1 Topología de red

Durante el proceso de definición del modelo se plantearon dos ubicaciones distintas para el IDS: integrado con el cortafuegos de entrada a la red o dentro de la red corporativa, trabajando en modo promiscuo.

En la primera opción, reflejada en la figura 2, todos los paquetes de entrada a la red pasarían por el IDS a la vez que por el cortafuegos, permitiendo al IDS colaborar de forma activa en la protección de la red, apoyando las decisiones del cortafuegos acerca de permitir o no en paso de paquetes.

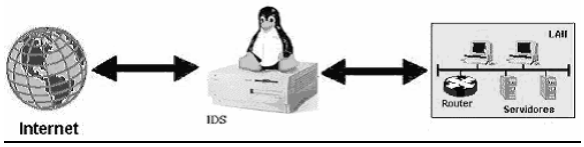


Fig. 2. Configuración de IDS integrado con cortafuegos

En la segunda opción propuesta (figura 3) se detectan tanto ataques externos a la red como ataques internos a ésta (los más peligrosos en general). Sin embargo el IDS pasa a tener un papel más pasivo que en el caso anterior, a no ser que se habilite un sistema de respuesta que permita la comunicación con elementos de protección activos de la red en tiempo real.

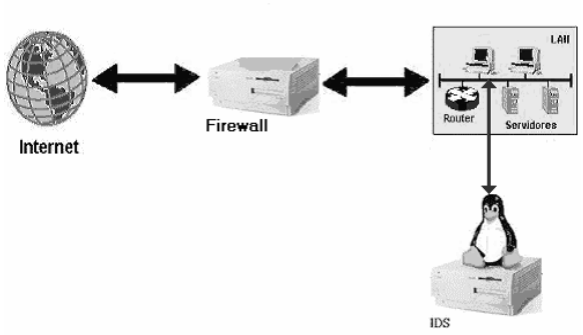


Fig. 3. Configuración de IDS interno a la red

La ubicación del IDS elegida ha sido la interna a la red trabajando en modo promiscuo. La elección se ha basado en el hecho de que el objetivo principal es realizar un análisis de los ataques sufridos por la red, de forma que se pueda apoyar a las decisiones que deba adoptar el administrador. Con este objetivo, un IDS interno a la red permite una mayor funcionalidad ya que puede analizar el tráfico interno y además no tiene grandes restricciones de tiempos de respuesta, favoreciendo los análisis basados en secuencias.

2.1.2 Tipo de IDS

Hay varios criterios de clasificación de los IDS en función de la característica a analizar. En este punto se usarán dos criterios distintos para definir el tipo de IDS: según las máquinas que entren en su ámbito de análisis y según el número de entidades en las que se distribuye el IDS.

Atendiendo al número de máquinas a las que monitoriza el IDS, se pueden diferenciar dos tipos: los basados en host (HIDS) que solo analizan una máquina, y los basados en red (NIDS) que analizan el intercambio de información entre todas las máquinas de una red.

Si el criterio es la distribución, los IDS pueden ser centralizados, si el proceso de análisis reside en una sola máquina, o distribuidos si realiza el análisis desde distintas máquinas coordinadas.

En este trabajo se ha optado por el uso de IDS basados en red (NIDS) y centralizados.

2.1.3 Tipo de Red Neuronal

El tipo de red elegido en este trabajo fue un asociador de patrones, implementada con una red multicapa feed-forward (perceptrón multicapa) con algoritmo de aprendizaje backpropagation (Regla Delta Generalizada). La robusted de este tipo de redes está demostrada en su aplicación en diferentes entornos, como detección de fraude en tarjeta VISA [8].

Lo único que falta para definir completamente la red es la topología de ésta. Sin embargo no es una cuestión sencilla ya que no hay criterios claros y objetivos para conocer a priori que una topología se comportará mejor que otra. En el desarrollo práctico se expondrán varias topologías ensayadas y sus resultados.

2.2 Arquitectura lógica

Una vez ubicado y definido el IDS y conocido el tipo de red neuronal a utilizar queda el diseño del motor de análisis, que será el elemento en el que se introducen las redes neuronales. Este diseño está en función de todos los elementos definidos anteriormente.

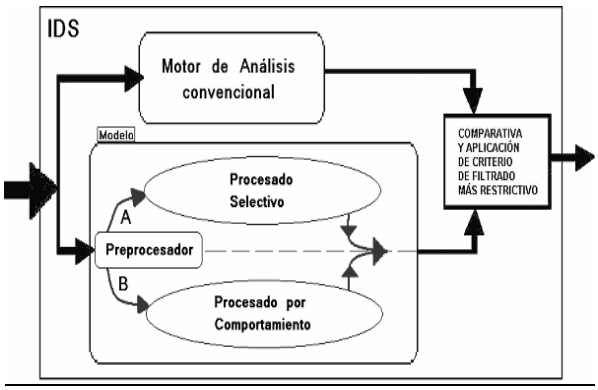


Fig. 4. Arquitectura de bloques del motor de análisis del IDS.

El motor de análisis que se plantea tiene dos mecanismos de procesamiento en paralelo. Por un lado está el motor de análisis del IDS elegido, basado en reglas normalmente, y otro módulo de procesamiento basado en redes neuronales. Ambos darán una salida indicando si existe o no una intrusión. Para unificar ambas salidas se introduce un bloque que permita la comparación y algún criterio de resolución de conflictos entre ambos.

2.2.1 Procesador basado en Redes Neuronales

Es el módulo encargado de la detección de intrusiones usando para ello redes neuronales. Como puede verse en la figura 4 está compuesto a su vez por dos procesadores en paralelo: uno denominado procesado selectivo y otro denominado procesado de comportamiento. Ambos están encargados de la detección de anomalías. Intentan analizar circunstancias distintas por lo que no son dos procesamientos para una misma detección.

El procesador selectivo realiza un análisis de cada paquete en tiempo real decidiendo si es o no un ataque.

El procesador de comportamiento evalúa las características del usuario que está generando eventos en la red. Para ello identifica el tráfico proveniente del mismo usuario y evaluará patrones de comportamiento de ese usuario, pudiendo detectar, por ejemplo, cuando hay un usuario y cuando un robot tras una serie de peticiones.

2.2.2 Procesador Selectivo

En el procesado selectivo intenta deducir si un determinado paquete de datos tiene estructura de ataque o no.

Abarcar en una misma red neuronal todos los posibles protocolos es inviable, ya que aumentaría el tamaño de la red de una forma inmanejable. Por este motivo se propone un procesamiento individual basado en protocolo.

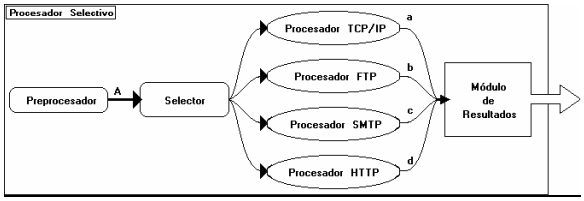


Fig. 5. Arquitectura del procesador selectivo.

Cada procesador de protocolo es una red neuronal distinta y entrenada de forma específica. El preprocesador prepara el paquete a analizar y lo entrega al selector. Este bloque analizará algunos parámetros del paquete y se lo entrega a los procesadores adecuados. Es posible que exista más de un procesador adecuado (por ejemplo un paquete WEB se puede analizar con un procesador IP, TCP y/o HTTP). Las respuestas de los procesadores serán agrupadas por un módulo de resultados, que avisará de las incidencias.

Cada procesador tiene una estructura como la mostrada en la figura 6.

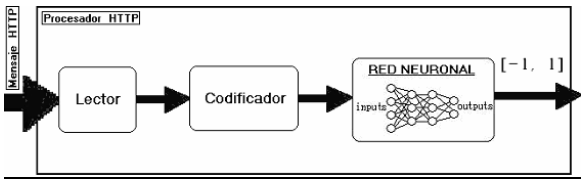


Fig. 6. Estructura del procesador de protocolo

En primer lugar hay un bloque denominado lector que se encarga de extraer la información relevante del fichero. La salida del lector irá al codificador que transforma la información del paquete a un formato entendible por la red neuronal. Finalmente a la salida del codificador estará la red neuronal que se encarga de determinar el grado de amenaza de la información recibida como entrada.

2.2.3 Procesador por Comportamiento

Este procesador se encarga de obtener, de entre todos los eventos de la red, información acerca del usuario que los genera. No intenta buscar fallos conocidos ni secuencias de ataques, sino modelar el comportamiento del usuario que genera los eventos de la red con el objetivo de ser capaz de identificar a los usuarios maliciosos.

Debido a que el comportamiento se suele definir a través de una sucesión de eventos es necesario que existan mecanismos de almacenamiento de la secuencia temporal de eventos o de composición de estos.

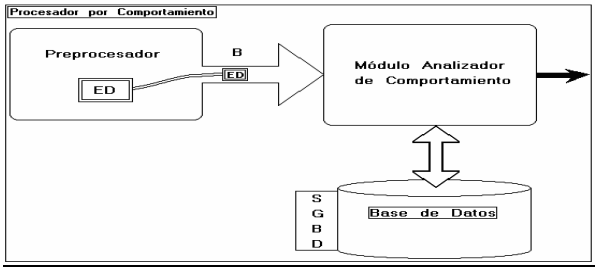


Fig. 7. Estructura del procesador de comportamiento.

El preprocesador usado es el mismo módulo que el del filtrado selectivo. Ahora recoge eventos de la red y forma una estructura de datos (ED) que pasa al módulo analizador de comportamiento. La estructura de datos tendrá la información suficiente para modelar el comportamiento del usuario. Para ello, en una primera fase, esta estructura debe tener los siguientes campos: origen y destino de datos (direcciones IP y espacios de aplicación), protocolos usados en la comunicación, sello de tiempo para conocer el momento en el que se recibió el evento y longitud de la información.

Una vez creada la estructura de datos se pasa al módulo analizador de comportamiento. En este módulo es donde se realizan las operaciones de análisis que tienen por objetivo detectar que un determinado usuario está realizando acciones ilícitas.

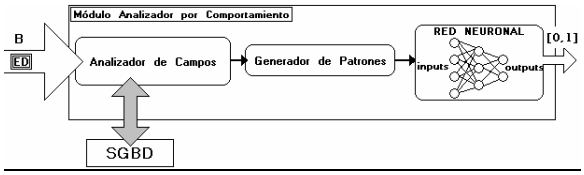


Fig. 8. Módulo analizador del comportamiento

El analizador de campos se encarga de recopilar la información de entrada a la red neuronal. A esta información se le ha denominado contexto de usuario. El contexto de usuario no solo contiene la estructura de datos (ED) que se acaba de recibir sino que también tiene toda la información almacenada en la base de datos y que está relacionada con la ED actual.

El generador de patrones se encarga de convertir el contexto de usuario recibido en un patrón legible por la red neuronal.

El patrón bien formado irá a la red neuronal que indicará si el usuario es o no un atacante.

3 Resultados

Para demostrar la viabilidad de la unión de los IDS y las redes neuronales se ha realizado una implementación práctica basada en el modelo propuesto en el apartado anterior.

En una primera fase del sistema desarrollado se ha probado el análisis selectivo, dejando para una segunda fase (actualmente en preparación) el análisis basado en comportamiento del atacante.

Dentro del análisis selectivo se ha centrado la investigación en el protocolo HTTP. La elección de este protocolo se ha debido a su gran difusión y a otros factores más técnicos como el hecho de que existen un gran número de ataques publicados en este entorno, lo que facilita el proceso de búsqueda de patrones.

Para realizar el estudio comparativo se han abordado, en paralelo, las siguientes líneas de trabajo:

- Integración del modelo en un IDS existente.
- Prueba del análisis selectivo basado en redes neuronales.

En el primer caso se ha modificado el motor de análisis de un IDS de libre distribución denominado Snort [9] para incluir el modelo basado en redes neuronales.

La segunda línea de trabajo es la más interesante para el objetivo de este trabajo ya que permite realizar una

validación del análisis selectivo sin la integración del IDS. A continuación se expondrán los puntos más relevantes del desarrollo realizado para llevar a cabo esta validación, así como los resultados obtenidos.

3.1 Obtención de patrones

Se deben encontrar patrones propios del protocolo a analizar (HTTP) correspondientes a comunicaciones normales y a intentos de ataques.

El cliente Web, accede a los servidores por medio de la URL. Hay muchos tipos de ataques Web. La mayoría de ellos se basan en la inserción de peticiones, mediante cadenas URL que producen algún tipo de daño: inyección de comandos, inyección de código SQL, Cross Site Scripting, etc.

Es fundamental para un correcto aprendizaje que la red tenga un gran número de ejemplos (cuantos más, mejor) de aquellos patrones que se pretende que aprenda a reconocer por generalización. El sitio web Security Focus (http://www.securityfocus.com) contiene un espacio dedicado a *bugs* de software (Sistemas Operativos, servidores, etc.). En este espacio existe un listado de ataques de todo tipo, del que se pudo extraer una lista de ataques basados en cadenas URL y servicios web y procesarlos.

Se obtuvieron un total de 850 ejemplos para el entrenamiento, y 60 para probar la red.[2]

3. 2 Codificación de patrones

El proceso de codificación es un proceso complicado ya que es necesario decidir, en primer lugar, qué información de los ejemplos es importante para darle a la red y que información no lo es. Una vez decidido, hay que pasarlo a un alfabeto entendible por la red.

Se han probado varias formas de realizar esta codificación, ya que condiciona la topología y el comportamiento de toda la red. La codificación finalmente utilizada se basa en la representación ASCII de los caracteres de la URL. Se dividen todos los caracteres entre 255 para que estén en el rango [0,1] y se aplica como entrada a la red neuronal. Previamente se sustituyen palabras no consideradas elementos peligrosos por símbolos predefinidos.[2]

Cadena			path=guestbook/print.php&id=1'									
Caracteres	@	=	@	/	@	.	@	&	@	=	1	'
Valores ASCII	255	61	255	47	255	46	255	38	255	61	49	39
Normalización	255/255	61/255	255/255	47/255	255/255	46/255	255/255	38/255	255/255	61/255	49/255	39/255
Resultado y entrada a la red	1	0.2392	1	0.1843	1	0.1803	1	0.149	1	0.2392	0.1921	0.1529

Fig. 9. Codificación normalizada

Se ha utilizado una segunda versión en la que no se normalizan los valores ASCII de los caracteres y se tratan a nivel de bit. Esto hace que aumente el número de entradas de la red .

Cadena	"<script>alert(document.cookie)</script>														
Caracteres	"	<	@	<	@	>	@	(	@	)	<	/	@	>	
Valores ASCII	34	62	255	60	255	62	255	40	255	41	60	47	255	62	
Normalización	34/255	62/255	255/255	60/255	255/255	62/255	255/255	40/255	255/255	41/255	60/255	47/255	255/255	62/255	
Resultado y entrada a la red	0.133	0.2431	1	0.2352	1	0.2341	1	0.1568	1	0.1607	0.2352	0.1843	1	0.2431	

Fig. 10. Codificación binaria

Para la codificación de los patrones a partir de las trazas originales se han desarrollado aplicaciones JAVA [3].

3.3 Análisis de topologías de red

Con el fin de crear y simular el comportamiento de la red neuronal, se ha utilizado el software JNNS (Java Neural Network Simulator) desarrollado en la Universidad de Stuttgart. Esta aplicación permite definir diferentes arquitecturas de redes neuronales, introducir patrones de entrada a través de ficheros, entrenar la red con estos patrones de entrada y, finalmente, realizar una simulación de su comportamiento.

En este trabajo se ha codificado además un software que permite generar una función matemática que representa el funcionamiento de la red y que es integrable en cualquier aplicación.

Como se comentó en apartados anteriores la definición de una topología, siguiendo parámetros objetivos, es algo complicado.

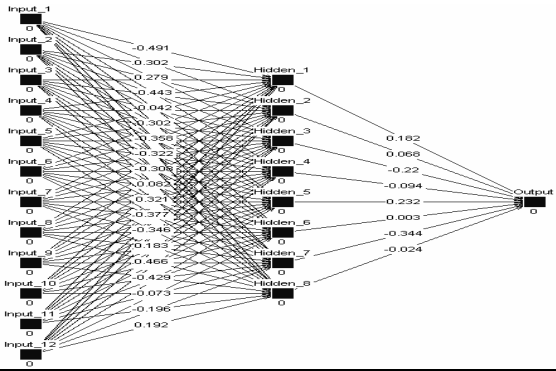


Fig. 11. Ejemplo de red neuronal.

Se han probado diversas arquitecturas, variables según el sistema de codificación usado y con un número de capas y unidades que es el que sigue:12-8-1, 12-9-5-1, y 10-7-4-1 con codificación normalizada y 96-20-1 y 96-30-15-4-1 con codificación binaria.

La clasificación escogida queda determinada por una salida bipolar con rango [-1, 1].

En la capa de entrada la función de salida era la unidad y en las capas ocultas la función *sigmoidal.*

3.5 Fase de entrenamiento

En la fase de entrenamiento, tras inicializar los pesos de manera aleatoria, comienza el aprendizaje de la red, que termina en el momento en que la función de error alcanza un mínimo.

Tal como se puede observar en la figura 12, el error referido en una de las topologías más significativas alcanza un mínimo estable tras unos 6000 ciclos de aprendizaje. A partir de este momento se está en condiciones de probar la eficacia de la red en el conjunto de patrones de test.

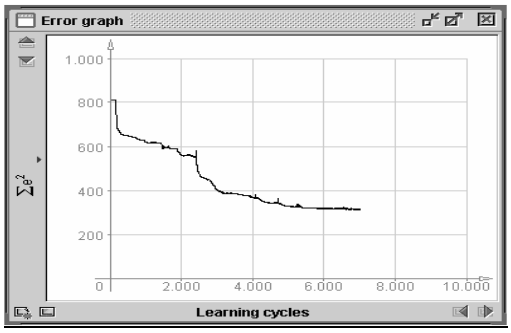


Fig. 12. Gráfica de error de la red 10-7-4-1

4 Discusión

La evaluación del paquete de *test* se llevó a cabo por medio de un programa Java, que en función de la salida obtenida, de la salida esperada y de un cierto umbral de error (escogido en función del grado de seguridad que se pretenda establecer) proporciona el porcentaje de aciertos de la red, que es el que se presenta a continuación.

TOPOLOGIA	CODIFICACIÓN	EFICIENCIA
12-8-1	Normalizada	63%
12-9-5-1	Normalizada	80%
10-7-4-1	Normalizada	88%
96-20-1	Binaria	68%
96-30-15-4-1	Binaria	95%

Fig. 13. Resultados obtenidos

De los fallos encontrados con la primera red, un 66.6% del total fueron falsos positivos (ataques considerados como conductas normales) y el resto correspondió a resultados ambiguos a interpretar por el gestor de seguridad. Estos datos demuestran que la red era capaz de detectar acciones normales de una forma más fiable, que por otro lado es lo que se pretende, pues siempre es preferible que se dejen pasar ataques a que se impida el paso a las acciones lícitas, lo cual sería propiamente un ataque.

Los resultados obtenidos en la segunda red, para un paquete de *test* de 60 patrones marcaron una efectividad del 88%, con 7 patrones erróneos de los cuales 4 correspondían a falsos positivos, 1 a un falso negativo y dos de ellos a resultados ambiguos (signo correcto pero por debajo del umbral). Una vez más, la red entrenada es menos restrictiva y detecta mejor las acciones normales.

En el caso de la topología 96-30-15-4-1 se obtuvo un acierto del 95%, lo que representa la mejor respuesta obtenida en todos los ensayos llevados a cabo. A priori no es previsible una tasa de acierto tan elevada, lo que sugiere el estudio del sistema con un conjunto más representativo de patrones.

Estos resultados suponen un punto de partida muy prometedor. Sin embargo sería necesario realizar un estudio acerca de la relación de los patrones para evaluar la dispersión de ataques que es capaz de manejar la red.

5 Conclusiones

El estudio que se ha llevado a cabo se ha centrado concretamente en el interés de aplicar las redes neuronales a los Sistemas de Detección de Intrusos (IDS) y se justifica por la capacidad de generalización de éstas. Esta capacidad de clasificar entradas no conocidas con anterioridad permite a los IDS descubrir intentos de intrusión no configurados previamente.

El modelo propuesto permite la integración de las redes neuronales en el proceso de análisis de los IDS. La integración trata de complementar el análisis tradicional realizado por el IDS para mejorar sus decisiones a la hora de detectar una intrusión. Adicionalmente el modelo da un nuevo enfoque para la detección de intrusos: no sólo analiza y monitoriza la secuencia de paquetes de información de la red usando técnicas propias de los sistemas inteligentes sino que trata de aprender el comportamiento de la persona que está accediendo a dicha red.

Para demostrar la viabilidad de esta integración se han probado varias configuraciones de redes neuronales aplicadas a la detección de intrusiones en secuencias de paquetes. Se han obtenido resultados variables según la topología elegida, llegando a más de un 90% de aciertos para los patrones de prueba disponibles, con un número bajo de falsos positivos (cuestión deseable para no filtrar acciones lícitas en la red). Sin embargo, es importante seguir evaluando arquitecturas de red y obteniendo nuevos patrones que sean significativos.

A partir del marco establecido en el presente trabajo se abren nuevas líneas de investigación:

- Diseño y evaluación de redes neuronales aplicadas a la clasificación del comportamiento de los usuarios de la red telemática, opción propuesta en el modelo teórico.

- Integración de las redes neuronales comentadas en el artículo en sistemas reales que permitan estudiar su comportamiento en entornos más ricos y complejos. Con esta integración se permite una evaluación más precisa de la red (ya que se compara con la salida del IDS real) y un aprendizaje más fino debido al gran número de patrones disponibles. Cuando la evaluación de resultados indicase que la red se comporta mejor que el análisis basado en reglas, se puede proceder a su sustitución.

- Implementación de sistemas de detección de intrusos distribuidos en varias máquinas, aprovechando el planteamiento intrísecamente distribuido de las redes neuronales.

Referencias

[1] J. Carracedo.“Seguridad en Redes Telemáticas” Ed. MacGraw-Hill. 2004.

[2] E. Ruiz, E. Torres Mejía .“Sistema inmunológico para la detección de intrusos a nivel de protocolo http” Proyecto de Grado presentado para optar al título de Ingeniero de Sistemas. Pontificia Universidad Javeriana (Bogotá). Mayo de 2003.

[3] D. Escarda Tejada y B. Jiménez Salmerón “Redes Neuronales en Sistemas de Detección de Intrusos”. Proyecto de Fin de Carrera. Diatel. EUITT. Madrid. Febrero de 2005.

[4] W. Fan, M. Miller, S. Stolfo, W. Lee, P. Chan “Using Artificial Anomalies to Detect Unknown and Known Network Intrusions”. *IEEE Intl. Conf. Data Mining*, pp. 123-130, 2001.

[5] M. Rash, A. Orebaugh, G.Clark. “IDS. Intrusion Prevention and Active Response: Deploying Network and Host IPS” Syngress, 2005.

[6] P. D.Wasserman. “Advanced Methods in Neural Computing”. Van Nostrand Reinhold, New York.1993

[7] L. V. Fausett. “Fundamentals of Neural Networks architectures. Algorithms. and applications”. Ed. Prentices Hall 1994

[8] J. R. Dorronsoro, F. Ginel, C. Sánchez, C. Santa Cruz. ``Neural Fraud Detection in Credit Card Operations" IEEE Trans. on Neural Networks 1--8, 1997.

[9] A. Baker, J. Beale, B. Caswell. “Snort 2.1 Intrusion Detection”. Syngress 2004

L-DPR: un esquema ligero de revocación de privilegios delegados

M. Francisca Hinarejos, Jordi Forné
Departamento de Ingeniería Telemática. Universidad Politécnica de Cataluña
ETSETB. C/ Jordi Girona 1-3.
08034 – Barcelona
Teléfono: 93 401 60 28 Fax: 93 401 59 81
E-mail: {mfcampos,jforne}@entel.upc.edu

Abstract*. The ubiquitous access allows users to establish communications anywhere and anytime. Different networks, topologies and technologies can be used: Wi-fi, UMTS, Bluetooth, Ethernet, etc. A challenging issue is to offer both authentication and authorization services based on certificates in this type of networks. The validation of the certificate's delegation path is a critical theme and requires several costly processes such as delegation path discovery or revocation status checking. The common solutions are not suitable in scenarios with limited resources: bandwidth, access to on-line servers, power energy, memory and computational capacity. We propose L-DPR, a lightweight revocation scheme that reduces the communication and computational overhead for the process of certificate status checking in delegation paths. Therefore the scheme is suitable for limited devices.*

1 Introducción

El acceso ubicuo permite que los usuarios puedan establecer comunicaciones en cualquier momento y en cualquier lugar. Diferentes redes, topologías y tecnologías pueden ser utilizadas: Wi-fi, UMTS, Bluetooth, etc. Tecnologías emergentes como MANET [1] permiten crear redes de forma espontánea, en las que no se requiere infraestructura. Sin embargo, proporcionar seguridad en este tipo de redes es más complejo que en las redes cableadas, debido principalmente a que: 1) son redes susceptibles a ataques, desde escuchas pasivas a interferencias activas, 2) existe una elevada probabilidad de que un nodo se rompa dentro de un largo periodo de tiempo, 3) la falta de soporte para infraestructura hace que los propios nodos deban gestionar la red, 4) la topología de la red es muy dinámica, debido a la movilidad de los nodos, 5) tanto la red como los dispositivos tiene ciertas limitaciones, como el ancho de banda o la batería respectivamente, etc.

La autenticación de los nodos es uno de los principales problemas en este tipo de redes, por lo que deben existir mecanismos que permitan la validación tanto de la identidad del nodo solicitando un servicio, como de la identidad del nodo ofreciendo el servicio. Soluciones habituales como Kerberos [2], X.509 [3] y PKIX [4] no son adecuadas para redes ad-hoc: 1) las soluciones centralizadas tienen problemas de escalabilidad, 2) el servidor de la CA (*Certification Authority*) está considerado como un único punto de fallo, 3) la movilidad de los nodos que forman la red hace difícil localizar al nodo que realiza las funciones de la CA, ya que las rutas cambian frecuentemente. De hecho, la disponibilidad del servicio es una de las principales dificultades en este tipo de escenarios. Por lo tanto, se necesitan nuevas soluciones adaptadas a estos nuevos entornos.

Existen propuestas orientadas a proporcionar servicios de autenticación en redes ad-hoc. Estas soluciones se pueden dividir en dos categorías: 1) en [5] y [6] se propone compartir la clave privada entre todos o un subconjunto de los nodos que forman la red. Estas soluciones están basadas en *threshold cryptography* [7]; 2) en [8] se propone crear una red de confianza de forma similar a PGP [9] donde los certificados son expedidos, almacenados y distribuidos por los propios usuarios, creando caminos de confianza entre ellos. Sin embargo, la revocación es todavía un tema de investigación no solventado por las soluciones comentadas.

El proceso de validación de certificados que debe llevarse a cabo en este tipo de redes requiere que se garanticen aspectos como: 1) el volumen de la información de revocación debe ser lo menor posible ya que tanto el ancho de banda como la capacidad de memoria de los dispositivos es limitada; 2) la información de revocación debe estar disponible en el instante de validación de las credenciales (mecanismos basados en esquemas on-line, como puede ser OCSP [10], no son adecuados; 3) el número de operaciones debe ser el menor posible debido a las limitaciones en la capacidad computacional de los dispositivos (en el proceso de validación es necesario verificar el estado de revocación de cada certificado en la cadena de certificados); y 4) la información de revocación debe estar actualizada para reducir la ventana de oportunidad donde los certificados revocados podrían ser utilizados para acceder a los servicios ofrecidos en la red.

Los aspectos arriba mencionados se agravan cuando

se utilizan certificados de atributo. Este hecho se debe a la diferencia de características entre los certificados de atributo y los certificados de identidad: los atributos asociados a una entidad (como puedan ser el rol de un usuario o sus privilegios asociados) cambian con mayor frecuencia en comparación con la clave pública del usuario, la probabilidad de revocación del certificado de atributo es mayor que el de identidad, los atributos delegados pueden llegar a crear largas cadenas de certificados, etc. Además, no existen soluciones propuestas para la revocación de certificados de atributos delegados para dispositivos móviles.

Nosotros proponemos L-DPR (*Lightweight-Delegated Privileges Revocation*), un esquema ligero de revocación basado en la identificación de certificados a través de palabras código. Esta codificación permite establecer tanto el camino de delegación en el proceso de validación (desde un determinado certificado hasta el certificado raíz), como los certificados expedidos a partir de un determinado certificado. Este hecho es posible ya que existe una dependencia jerárquica de los certificados de atributo que permite su representación a través de un árbol binario. Este esquema permite reducir tanto el volumen de información de revocación como la carga del proceso de validación en cadenas de delegación. Estas características hacen que el esquema de revocación propuesto sea adecuado para dispositivos móviles con recursos limitados.

El resto del artículo se estructura de la siguiente manera. En el apartado 2 se presenta una clasificación de la revocación de privilegios existente en la literatura, junto con algunos de los mecanismos de obtención de la información de revocación. Los problemas relacionados con la revocación en caminos de delegación se exponen en el apartado 3. El funcionamiento del esquema propuesto se explica en el apartado 4. Por último, se presentan las conclusiones en el apartado 5.

2 Revocación de privilegios

Un certificado conteniendo los privilegios de un usuario, se debe revocar cuando alguno de los privilegios asociados deja de ser válido. Existen diferentes motivos para revocar un certificado, como pueda ser el uso fraudulento de los privilegios por parte del usuario. Cuando un certificado de atributo está revocado, el resto de entidades deben poder verificar que se ha revocado el certificado.

En este apartado se explican dos procesos relacionados con la revocación de privilegios: 1) los pasos para revocar un privilegio (o el certificado que contiene dicho privilegio), y 2) los mecanismos utilizados para verificar el estado de revocación del certificado.

2.1 Políticas de revocación

En determinados escenarios se necesita delegar privilegios o derechos de usuario, de una entidad a otra. Este proceso debe ser realizado por entidades autorizadas. La revocación de los privilegios delegados debería ser realizada por entidades autorizadas, tales como: el poseedor del privilegio, la entidad que delegó el privilegio u otra entidad autorizada. Otro aspecto a tener en cuenta es cómo afecta la revocación de un privilegio sobre el resto de privilegios, ya sea directa o indirectamente.

Para poder estudiar los diferentes escenarios que podrían presentarse a la hora de revocar privilegios, los autores en [11] diferencian tres dimensiones:

- *Resilence*. Hace referencia a la persistencia de la revocación en el tiempo:
  1) La revocación por eliminación del privilegio, la cual tendría un carácter temporal, ya que se podría volver a asignar el privilegio en un tiempo posterior.
  2) La revocación vía la asignación de un privilegio negativo, el cual anula al privilegio positivo inicialmente asignado, hasta que el privilegio negativo sea eliminado.
- *Propagation*. A diferencia del caso anterior en el cual la revocación tenía carácter temporal, en este caso, la revocación tiene carácter espacial, es decir, a qué usuarios afecta la revocación de un determinado privilegio:
  1) Local: si sólo afecta al usuario al cual se le revoca el permiso.
  2) Global: si la revocación es en cascada, es decir, afecta al usuario y al mismo permiso concedido a otras entidades por el usuario al cual se le revoca inicialmente el permiso.
- *Dominance*. A un usuario se le revoca un permiso concedido por una determinada entidad, sin embargo, todavía tiene permisos procedentes de otras entidades. Dentro de este modelo se pueden diferenciar:
  1) *Strong revocation*: Si las entidades son independientes, los permisos concedidos no tienen porqué verse afectados. Si las entidades recibieron sus permisos a través de una cadena en la cual se encuentra la entidad que desea revocar el permiso, ésta puede revocar dichas asignaciones.
  2) *Weak revocation*: si las entidades son independientes a la entidad que revoca el permiso, sólo se ve afectado el permiso concedido directamente por la entidad que lo revoca.

Esta clasificación no deja de ser conceptual, ya que si se aplica a la revocación de privilegios utilizando certificados de atributo X.509 [3], ésta puede llegar a

ser inviable. Por lo tanto, nosotros partimos de la hipótesis que cuando se revoca el certificado de atributo de un usuario, éste deja de ser válido indefinidamente, y tiene un efecto en cascada sobre los certificados de atributo que se expidieron a partir de él.

2.2 Verificación del estado de revocación

Una vez un certificado (o el privilegio contenido en el certificado) se ha revocado, debe existir algún mecanismo que permita verificar si un certificado es o no válido. Existen diversos mecanismos de revocación que permiten obtener la información sobre los certificados que están revocados. La distribución de la información de revocación se puede llevar a cabo a través de dos caminos:

- Offline: las entidades utilizadas para la distribución de la información del estado de revocación son no TTPs (*Trust Third Party*). En este caso, el expedidor pre-computa los datos del estado de revocación y los distribuye entre los repositorios, como puede ser CRL [3]. Por lo tanto, la confianza sobre la entidad que distribuye la información disminuye, pero el volumen de información contenida en la CRL aumenta con el número de certificados revocados.
- Online: las entidades utilizadas para la distribución de la información del estado de revocación son TTP. En este caso, el servidor ofreciendo el servicio proporciona la evidencia criptográfica de los datos del estado de revocación de los certificados, como puede ser OCSP [10]. De esta manera se consigue reducir el ancho de banda necesario para la distribución de la información de revocación aumentando la confianza en la entidad que proporciona la información.

Para la verificación de la validez de un certificado, se debe verificar que ningún certificado involucrado en la cadena está revocado. Para poder comprobar el estado de revocación de cada uno, primero se debe conocer el identificador de los certificados en la cadena, y a continuación solicitar la información de revocación de cada uno de ellos. Este procedimiento se puede agravar cuando se permite la delegación de privilegios, ya que el número de certificados en la cadena puede ser elevado.

En el apartado 3 se explica con mayor detalle los problemas relacionados con la revocación en los caminos de delegación.

3 Aspectos de la revocación en caminos de delegación

En algunos escenarios los atributos o privilegios de una entidad pueden ser delegados a otras entidades. Estas delegaciones pueden ser representadas por estructuras jerárquicas. La Fig. 1 representa un ejemplo de una posible relación jerárquica de la delegación de privilegios entre diferentes entidades. Supongamos que la entidad *A* posee un conjunto de cuatro privilegios {1,2,3,4}, de los cuales delega un subconjunto de sus privilegios a *B* {1,2} y el subconjunto de privilegios {3,4} a la entidad *C*. De forma similar, la entidad *C* delega los privilegios {1,4} a la entidad *D*. Y finalmente, la entidad *D* delega el privilegio {4} a la entidad *E*. La delegación sólo es posible cuando las entidades poseen autoridad para delegar privilegios. De hecho, la longitud del camino de delegación puede ser limitada por la entidad que es fuente de autoridad de los privilegios. Es decir, si la entidad *A*, la cual en el ejemplo de la Fig. 1 es la fuente de autoridad del conjunto de privilegios {1,2,3,4}, fija la longitud máxima del camino de delegación a 2, la entidad *E* no puede delegar el privilegio {4} a ninguna otra entidad.

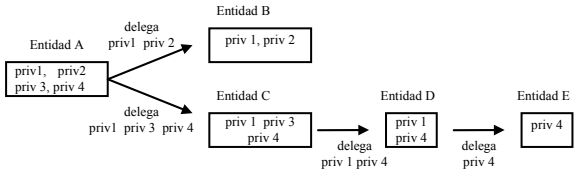


Fig.1 Proceso de delegación de privilegios. Relación entre las entidades y los privilegios delegados.

A veces es necesario revocar un privilegio a una determinada entidad debido, por ejemplo, a la utilización fraudulenta de los privilegios por parte del usuario que los posee. Si *A* revoca el privilegio {4} a *C*, el mismo privilegio, concedido por *C* a otras entidades, debe dejar de ser válido. Por lo tanto, el privilegio 4 concedido tanto a la entidad *D* como a la entidad *E* también deben dejar de ser válidos.

Supongamos ahora que los privilegios están contenidos en certificados digitales, y que se revoca el privilegio {4} de la entidad *C* (ver Fig. 2). Para poder verificar si la entidad *E* posee un determinado privilegio, un verificador de privilegios debe obtener los certificados que forman la cadena de delegación, desde el certificado de la entidad *A* hasta el certificado de la entidad *E* {*A,C,D,E*} (este proceso se conoce como *path discovery*). A continuación, el verificador de privilegios debe chequear si cada uno de los certificados en la cadena está revocado. El descubrimiento de cadenas de certificados es un proceso costoso [12,13] y debe llevarse a cabo incluso si alguno de los certificados en la cadena está revocado. Este hecho es debido a que la entidad que intenta verificar la cadena de certificados, a priori no conoce el identificador de los certificados que la forman. Por lo tanto, dicha entidad no puede solicitar información sobre el estado de revocación de los certificados que forman la cadena de delegación.

Una solución sencilla para el problema anterior sería incluir el identificador de los certificados del camino de delegación en cada certificado expedido (desde el certificado raíz, hasta el expedidor del certificado

objeto). Esta información podría ser incluida en el certificado como una extensión[1] del certificado a expedir [3]. Por ejemplo, en la Fig. 2, el certificado de la entidad *E* debería incluir el identificador de los certificados de las entidades *A*, *C* y *D*. En este caso, cuando una entidad intenta validar el certificado de la entidad *E*, debe obtener los identificadores incluidos en una extensión del certificado de *E*, y verificar el estado de revocación de cada certificado identificado. Si cualquiera de los certificados que forman la cadena está revocado, no sería necesario llevar a cabo el proceso de *path discovery*. Sin embargo, sí se debe verificar el estado de revocación de cada uno de los certificados.

Una alternativa podría ser actualizar las dependencias entre los certificados cuando se produce la revocación. En otras palabras, cuando se revoca un certificado, entonces también se revocan los certificados relacionados jerárquicamente con el certificado que se revoca. En la Fig. 2, cuando se revoca el certificado $Cert_C$, los certificados $Cert_D$ y $Cert_E$ también se revocan. La solución de incluir los identificadores del camino de delegación podría no ser adecuada ya que la información incluida es referente a los certificados desde el certificado raíz hasta el expedidor del certificado objeto. En este caso, se necesita la información relativa a los certificados expedidos a partir de un determinado certificado, pero adquirir este conocimiento no es una tarea sencilla.

Para solventar este inconveniente, sería necesario realizar un seguimiento del camino de delegación. De esta manera, cuando los privilegios de la entidad *C* contenidos en el certificado $Cert_C$ dejaran de ser válidos, los certificados $Cert_C$, $Cert_D$ y $Cert_F$ también se marcarían como inválidos. En este caso, sólo es necesario validar el estado de revocación del certificado objeto, por lo que el estado de revocación de $Cert_F$ reflejaría el estado de revocación del camino completo de delegación. Este procedimiento permite proporcionar los identificadores del camino de delegación, así como ayudar en el proceso de *path discovery*. Sin embargo, se requiere una entidad de confianza para actualizar el árbol jerárquico representando las dependencias entre los certificados. En [19] se menciona la necesidad de llevar a cabo el seguimiento de cadenas sobre cadenas de certificados de clave pública concretas, pero únicamente dentro de un ámbito local. Es decir, la información de revocación sólo se utiliza bajo el dominio de un servidor y durante un periodo de tiempo limitado.

Este tipo de escenarios son en los que nos basaremos, es decir, se llevará a cabo una política de revocación en cascada, por lo que una vez revocado un certificado, se deben revocar sus dependencias. Esta es una primera aproximación y esperamos contemplar otro tipo de situaciones en las que la revocación de un certificado no implique la revocación de todos los certificados conteniendo los privilegios delegados. Esta situación podría suponer la expedición de nuevos certificados a todos los usuarios por debajo del certificado revocado, lo que conllevaría una elevada carga de gestión.

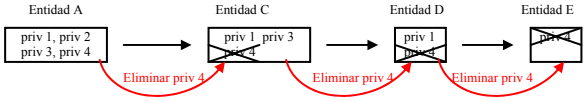


Fig.2 Proceso de cancelación de privilegios. Proceso de revocación de privilegios delegados a entidades con dependencias jerárquicas.

4 Esquema de revocación propuesto

Nosotros proponemos L-DPR (*Lightweight-Delegated Privileges Revocation*) un esquema de revocación basado en la codificación de los identificadores de los certificados a través de palabras-código adecuadas. Como se explica en este apartado, esta codificación permite llevar a cabo una política de revocación en cascada. Esta sencilla técnica permite obtener las siguientes ventajas: 1) proceso simple de descubrimiento de caminos de delegación, y 2) reducir la complejidad en el proceso de verificación del estado de revocación de los certificados que forman el camino de delegación. El último punto se consigue evitando tener que verificar el estado de cada certificado en el camino de delegación. Los procedimientos para lograr estos objetivos se explican a continuación.

4.1 Codificación del árbol

Como ya se ha comentado, las dependencias entre los certificados pueden ser representadas a través de un árbol jerárquico[2]. La Fig.3 representa las dependencias de los certificados en un árbol lógico: 1) cada nodo representa un certificado de atributo y cada rama, desde el nodo raíz hasta el nodo que representa el certificado objeto, representa el camino de delegación; 2) si se revoca el certificado representado por el nodo 2, los nodos 3, 4, 5, 6 y 7 deben dejar de ser válidos debido a las dependencias entre los nodos.

Si a cada nodo se le asigna un identificador diferente, el certificado puede ser identificado de forma unívoca dentro del árbol. Además, si el identificador se genera de forma adecuada, a partir de la información contenida en el nodo padre, se consigue un esquema

[1] El campo *extension* en un certificado [3] permite agregar información adicional no contemplada en los campos obligatorios.

[2] Las relaciones en el proceso de delegación pueden dar lugar a grafos, sin embargo, este enfoque está fuera del estudio de este artículo.

que permite tanto representar las dependencias de los certificados como identificar los certificados de forma unívoca.

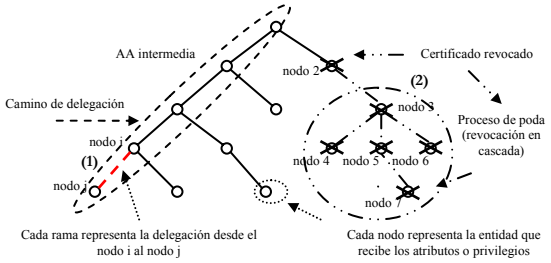


Fig.3 Representación lógica de las dependencias entre los certificados de privilegios.

Ahora se debe escoger una técnica de codificación adecuada para lograr las características antes mencionadas, es decir:

- El identificador del certificado debe ser único dentro del dominio de cada autoridad de atributos.
- El identificador debe proporcionar información para poder seleccionar los certificados involucrados en el camino de delegación (nosotros denotamos a este hecho como *down-up knowledge*).
- El identificador debe proporcionar información para poder discernir cuáles son los certificados expedidos a partir de los atributos de un determinado certificado (*up-down knowledge*). De esta manera, se puede conocer de forma implícita la revocación de los certificados dependientes, sin tener que obtener los identificadores de tales certificados.

En este artículo adoptamos la aproximación realizada por los autores en [15] para la construcción de un árbol binario a partir de un árbol multi-camino (la misma adoptada en [16] para llevar a cabo el proceso de *path discovery*). En este caso, cada palabra-código de un nodo hijo se genera como la concatenación de la palabra-código asignada al nodo padre y una nueva palabra-código generada, es decir:

- Al primer nodo hijo se le asigna una palabra-código generada por la concatenación de la palabra-código del padre y un 0.
- La palabra-código del resto de hijos en el mismo nivel, se genera como la concatenación de la palabra-código del primer hijo y un número de unos igual a *n-1*, donde *n* se corresponde con el número de hijo bajo el dominio directo del nodo padre.

Esta codificación permite que de forma directa se pueda transformar un árbol multi-camino en un árbol binario. La Fig.4 representa un posible procedimiento lógico para generar una palabra-código como se ha descrito antes.

Este tipo de codificación permite proporcionar una palabra-código diferente por cada nodo del árbol a ser generado, consiguiendo que el identificador de cada certificado sea único. De esta manera, se consigue obtener un conocimiento implícito del camino de delegación, desde el nodo raíz hasta el certificado objeto (*down-up knowledge*) [16]. Aunque la característica más importante para nuestro esquema de revocación es que permite determinar cuáles son los certificados dependientes a un determinado certificado objeto, de una forma implícita (*up-down knowledge*). Por ejemplo, en la Fig. 5 la palabra-código D→X10 permite obtener el identificador de los certificados que forman la cadena de delegación, C(X1) y A(X), ya que la palabra-código D contiene la información necesaria. Por otra parte, cualquier certificado que contenga como prefijo la palabra-código "X10", depende del certificado con palabra-código "X10". De esta forma, tanto H1(X100) como H2(X1001) son certificados cuyos privilegios han sido delegados a partir de los privilegios contenidos en D(X10).

```
function codeWordGeneration ( cw_parent , pos_child )
{
    cw_child = cw_parent || 0
    // if the first child
    if ( pos_child == 0) then
            return ( cw_child );
    for (int i=0;i< pos_child ;i++)
    {
            cw_child = cw_child || 1 ;
    }
    return ( cw_child );
}
```

La primera palabra-código se genera como la concatenación de la palabra-código del padre y un 0

El resto de palabras-código se generan como la concatenación de la palabra-código del primer hijo y n-1 unos, donde n representa el número de hijo bajo el dominio directo de un nodo padre

Fig. 4. Pseudo-código del algoritmo para generar las palabras-código representando cada certificado de atributo en un nivel.

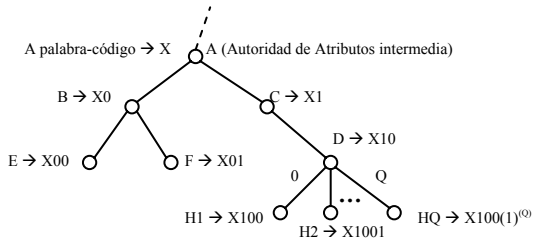


Fig. 5. Árbol de codificación representando los certificados de atributo.

4.2 Palabras-código en certificados de atributo

Los certificados de atributo deben incluir su identificador, que en nuestro esquema es una palabra-código. Bajo la misma autoridad de atributos cada certificado se diferencia del resto por un número de serie. Por lo tanto, el campo que contiene el número de serie, es el más adecuado para contener la palabra-código.

La longitud de la palabra-código puede ser elevada si el número de certificados expedidos por la misma autoridad es elevado. Para reducir el volumen de información contenido en un único campo del

certificado, se podría dividir la información para ser transportada en diferentes campos dentro del certificado.

La palabra-código identificando un certificado se puede dividir en dos partes (ver Fig. 6): 1) la palabra-código del certificado padre, el cual se corresponde al prefijo, y que denotaremos como *prefix-code*; y 2) el código que identifica a un certificado bajo una misma autoridad padre, que denotaremos como *child-code*. Un certificado digital contiene un campo para identificar al poseedor del certificado, y otro campo para identificar al expedidor del mismo. De esta manera, el campo para identificar al expedidor se puede utilizar para transportar el *prefix-code*, y el campo del número de serie puede ser utilizado para contener el *child-code*. Esta división permite incrementar el número de certificados expedidos en cada nivel.

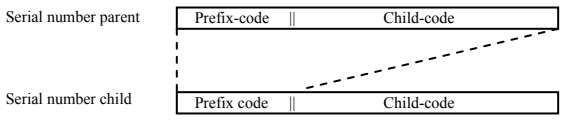


Fig. 6. Codificación del número de serie en certificados de atributo.

4.3 Compresión del identificador del certificado

En el apartado 4.2 se ha comentado que una palabra-código puede tener una longitud elevada si el número de certificados expedidos directamente por una misma autoridad es elevado, o si la profundidad del árbol multi-camino es elevada. En otras palabras, el número máximo de bits necesarios para representar una palabra-código es $q \log_q n$ donde q representa el número máximo de certificados expedidos por la misma autoridad de atributo, y $\log_q n$ representa la profundidad del árbol. Para reducir el volumen de la información del identificador de un certificado, proponemos utilizar un algoritmo de codificación de fuente.

La palabra-código generada en el apartado 4.1 sigue una estructura bien conocida. Esta estructura puede ser generalizada para cada nivel del árbol, como una secuencia $\{01...1\} = \{01^q\}$ donde 1^q representa a una secuencia de q unos. *Run Length Encoding* (RLE) o codificación de ráfagas, es una algoritmo de compresión de datos simple y bien conocido, el cual es muy útil para comprimir este tipo de secuencias [17]. RLE está basado en la idea de reemplazar una secuencia larga de un mismo símbolo por una secuencia más corta que la represente. La secuencia de longitud l de un símbolo repetido 's', se reemplaza por una secuencia más corta, normalmente, la secuencia por la que se reemplaza contiene: uno o más símbolos de 's', la información sobre la longitud, y a veces, el símbolo de *escape*. Por ejemplo, si una autoridad de atributos expide 10000 certificados, la última palabra-código generada contendrá unos 10000 bits (1250 bytes aproximadamente), más los bits de la palabra-código que representa el certificado de la autoridad de atributos. Esta información contiene un alto grado de redundancia y reduce la viabilidad del esquema de revocación. Sin embargo, si el *child-code* se genera como un valor entero indicando el número de unos, para el caso anterior, sólo serían necesarios 14 bits para representar al mismo certificado (aproximadamente 2 bytes). Por lo tanto, se consigue reducir el volumen de información a transportar en el certificado.

4.4 Generación del árbol de revocación

El proceso de revocación conlleva principalmente 2 subprocesos: 1) la petición de revocación de un certificado [14], y 2) el mecanismo para obtener la información de revocación [1, 10, 18]. La petición de revocación de un certificado se puede llevar a cabo a través de los procedimientos ya existentes, ya que la información principal contenida en la petición es el identificador del certificado, que en nuestro esquema es una palabra-código. Este hecho permite reutilizar los protocolos existentes para la petición de revocación de certificados. Del mismo modo, alguno de los mecanismos de obtención de la información de revocación también podría ser reutilizado, como pueda ser OCSP.

Definimos como *revocation tree*, el árbol que contiene la información de revocación. Cuando se recibe una petición de revocación, se debe actualizar el árbol de revocación de acuerdo a la política dada. En primer lugar, se debe verificar si la palabra-código a ser revocada es un nodo interno del árbol de revocación[3]. Si es un nodo interno, se deben podar todas las ramas que parten del nodo que representa a la palabra-código, si así lo especifica la política de revocación. Esta característica permite reducir el volumen de la información de revocación cuando la revocación involucra más de un certificado. Por otra parte, si la palabra-código no existe, se debe crear una nueva entrada en el árbol.

La Fig. 7 (a) representa un árbol de revocación con tres certificados revocados. En este caso, se revoca el certificado identificado por la palabra código X1. X1 es prefijo tanto de la palabra-código X100 como de X1011, por lo tanto, la palabra-código X1 representa de forma implícita tanto a X100 como a X1001. De esta forma, sólo es necesario representar en el árbol la palabra-código X1, reduciendo el volumen de la

[3] El nodo hoja en el árbol contiene la información completa de los certificados revocados, pero un nodo intermedio no representa a un certificado revocado, a diferencia del árbol de codificación (ver apartado 4.1) donde cada nodo representa un certificado expedido.

información de revocación a ser almacenada (ver Fig. 7 (b)), y en su caso, a ser transmitida.

Cada vez que una entidad debe verificar el estado de un certificado, sólo necesita verificar el estado de revocación del certificado objeto. En otras palabras, la entidad intentando validar un certificado, no necesita verificar el estado de revocación de cada certificado que forma la cadena de delegación, a diferencia de las soluciones existentes en las que se debe obtener el camino completo de certificados, y validar el estado de cada certificado.

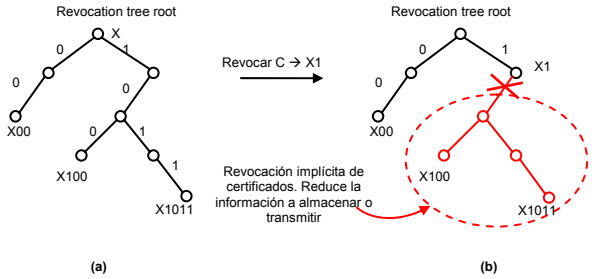


Fig. 7. Representación del árbol de revocación. (a) árbol de certificados revocados. (b) Nuevas revocaciones permiten reducir el volumen de información del árbol de revocación.

4.5 Verificación del estado de revocación

El proceso de verificación del estado de revocación es un proceso sencillo. Por ejemplo, si se quiere verificar el estado del certificado identificado por X1011 sobre el árbol representado por la Fig. 7 (b), se realizan los siguientes pasos (ver Fig. 8):

- Obtener cada dígito binario de la palabra código, empezando desde el dígito más significativo (izquierda) hasta el menos significativo (derecha).
- El primer dígito es un "1", el cual se corresponde con la rama derecha del árbol.
- El nodo X1 es tanto una hoja (es un certificado revocado) como el prefijo de la palabra código X1011. Por lo tanto, el certificado X1011 está revocado de forma implícita. En este punto, el proceso de verificación del estado de revocación del certificado finaliza.

5 Conclusiones

La validación de los caminos de delegación es un aspecto crítico en infraestructuras basadas en certificados digitales, ya que se requieren llevar a cabo algunos procesos costosos tales como el descubrimiento de cadenas de certificados, y el proceso de verificación del estado de revocación de cada certificado en el camino de delegación. Las soluciones habituales no son adecuadas en escenarios con recursos limitados tales como ancho de banda, batería y dispositivos limitados tanto en memoria como en capacidad computacional.

El esquema de revocación propuesto permite reducir el conocimiento necesario de los certificados que forman el camino de delegación en el momento de la validación. Para poder saber si cualquier certificado de la cadena de delegación está revocado, sólo es necesario verificar el estado de revocación del último certificado en el camino de delegación (el primer certificado es el certificado de la fuente de autoridad). Este hecho se consigue gracias a la codificación especial del identificador del certificado a través de palabras-código. De esta manera, el esquema de revocación propuesto permite reducir tanto la carga computacional como el ancho de banda necesario. Estas características hacen que L-DPR sea adecuado en escenarios limitados como puedan ser las redes ad-hoc con acceso intermitente a sistemas on-line.

```
function verifyCWRevocation ( cw_certificate )
{
    cw_i = cw_certificate ;
    l_cw = |cw_i| ;
    pointer = tree.root;
    for (int i=0;i< l_cw ;i++)
    {
        if pointer.left.bit != cw_i(i) || not_exist then
        {
            if pointer.right.bit != cw_i(i) || not_exist then
                return(not revoked);
            else if pointer.isleaf then
                return(revoked);
            else
                pointer = pointer.right;
        }
        else
        {
            if pointer.isleaf then
                return(revoked);
            pointer = pointer.left;
        }
    }
}
```

Fig. 8. Pseudo-código representando el algoritmo que permite verificar si un certificado de atributo ha sido revocado, siguiendo los pasos explicados en el apartado 4.5.

Agradecimientos

Este trabajo ha sido realizado con el soporte del Departament d'Universitats, Recerca i Societat de la Informació. Trabajo parcialmente subvencionado por el Ministerio de Ciencia y Tecnología dentro del proyecto ARPA TIC2003-08184-C02-02.

Referencias

[1] RFC 2501. S. Corson, University of Maryland, J. Macker, Naval Research Laboratory. Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations. January 1999.

[2] RFC 1510. The Kerberos Network Authentication Service v5, September 1993.

[3] ITU-T Recommendation X.509, Information technology – Open Systems Interconnection – The Directory: Public Key and Attribute Certificate Frameworks. 2000.

[4] PKIX: http://www.ietf.org/html.charters/pkix-charter.html.

[5] J. Kong, P. Zerfos, H. Luo, S. Lu and L. Zhang. Providing robust and ubiquitous security support for manet. Proceedings of the 9th IEEE International Conference on Network Protocols (ICNP), pp. 251-260, 2001.

[6] L. Zhou and Z.J. Haas. Securing ad hoc networks. IEEE Network Magazine, vol 13, no 6, pp. 24-30, November/December 1999

[7] Desmedt Y. and Frankel Y. Threshold cryptosystems. In Advances in Cryptology—Crypto'89, the Ninth Annual International Cryptology Conference, Proceedings, Lecture Notes in Computer Science, vol. 435., G. Brassard, Ed., Springer-Verlag, Berlin, 307–315.

[8] Srdjan Capkun, Levente Buttyan and Jean-Pierre Hubaux . Self-Organized Public-Key Management for Mobile Ad Hoc Networks IEEE Transactions on Mobile Computing, Vol. 2, No. 1, Jan-Mar 2003.

[9] Pretty Good Privacy, http://www.pgpi.org.

[10] RFC 2560. Myers M., Ankney R., Malpani A., Galperin S., and C. Adams, X.509 Internet Public Key Infrastructure – Online Certificate Status Protocol – OCSP. June 1999.

[11] Hagstrom A.; Jajodia S.; Parisi-Presicce F.; Wijesekera D., Revocations –a classification. Computer Security Foundations Workshop, 2001. Proceedings. 14th IEEE, 11-13 June 2001. Page(s): 44 -58.

[12] RFC 3379. D. Pinkas, Bull, R. Housley, RSA Laboratorios, Delegated Path Validation and Delegated Path Discovery Protocol Requirements. September 2002.

[13] Steve Lloyd, PKI Forum, Understanding Certification Path Construction. White Paper. September 2002.

[14] RFC 2510. C. Adams, Entrust Technologies, S. Farell, SSE, Internet X.509 Public Key Infrastructure Certificate Management Protocols. March 1999.

[15] J. Stasko, J. Vitter, Pairing Heaps: Experiments and Analysis, In Communications of the ACM, March 1987.

[16] He Huang, Shyhtsun Felix Wu, An approach to certificate path discovery in mobile Ad Hoc networks. Proceedings of the 1st ACM workshop on Security of ad hoc and sensor networks, October 2003, pp. 41-52, ISBN:1-58113-783-4.

[17] M. Nelson, The Data compression book. M&T Books, New York, NY 1995. ISBN 1-55851-434-1.

[18] S. Micali, Efficient Certificate Revocation, Massachusetts Institute of Technology, Cambridge, MA, 1996.

[19] Popescu, B.C.; Crispo, B.; Tanenbaum, A.S: A certificate revocation scheme for a large-scale highly replicated distributed system. Computers and Communication, 2003. (ISCC 2003). Proceedings. Eighth IEEE International Symposium on , 2003 Page(s): 225 -231.

Protección contra el Spam utilizando Desafíos “a priori”

Rodrigo Román[1], Javier López[1], Jianying Zhou[2]
[1] E.T.S. Ingeniería Informática, Universidad de Málaga, 29071, Málaga, España
[2] Institute for Infocomm Research, 21 Heng Mui Keng Terrace, Singapore 119613
roman@lcc.uma.es, jlm@lcc.uma.es, jyzhou@i2r.a-star.edu.sg

Abstract*. Spam is considered to be one of the biggest problems in messaging systems. In the area of email Spam, A high number of anti-spam schemes have been proposed and deployed, but the problem has yet been well addressed. In this paper, we introduce a new scheme, called pre-challenge scheme, which avoids problems that exists in other schemes such as delay of service and denial of service. Some new mechanisms are employed to reach a good balance between security against Spam and convenience to email users. In addition, our scheme can be used for protecting other types of messaging systems, such as Instant Messaging (IM) and Blogs, against Spam.*

1 Introducción

El *Spam* (Mensajes electrónicos no solicitados) es considerado como uno de los mayores problemas de los mecanismos de mensajería sobre Internet. Mediante una inversión mínima, es posible tanto inundar con propaganda no deseada a cualquier usuario como aumentar la visibilidad de una página web en motores de búsqueda de forma fraudulenta.

El correo electrónico es el sistema más afectado por este problema. El volumen de *Spam* (“correo basura” en este contexto) recibido por cualquier usuario puede ser tal que el tiempo que se necesita para separar los correos importantes de los correos basura llega a ser prohibitivo. Es por esto por lo que existen multitud de soluciones que tratan de evitar, con mayor o menor éxito, que los correos basura alcancen los buzones de los usuarios.

Específicamente, existe una solución denominada “Desafío/Respuesta” (“Challenge/Response”), en la que se obliga a los emisores a resolver un desafío antes de poder acceder al buzón del receptor. Esta solución posee varios problemas que dificultan su uso, tales como la introducción de un tiempo de espera antes de que el receptor reciba realmente el mensaje (ya que el desafío se envía al emisor cuando el sistema que maneja los mensajes del receptor recibe su correo), la posibilidad de realizar ataques de denegación de servicio (si la dirección del emisor de un mensaje está falsificada), y otros problemas como el manejo de listas de correo y de mensajes de error.

En este artículo proponemos una nueva solución, denominada desafíos “a priori”, la cual posee todos los beneficios de los mecanismos de “desafío/respuesta” aplicados al correo electrónico sin ninguno de sus inconvenientes, y que además puede aplicarse a otros problemas de mensajes no solicitados, tales como la mensajería instantánea y los Blogs. Esta solución también incluye mecanismos que pueden aplicarse de forma independiente a los gestores de correo actuales, como el mecanismo de manejo de mensajes de error o la “lista de alerta”.

Este artículo se organiza de la siguiente forma: En la sección 2, se introducen cuales son los principales problemas en los sistemas de mensajería. En la sección 3 se analiza el trabajo previamente realizado para proteger a la infraestructura del correo electrónico contra el *Spam*. En la sección 4 se introduce la nueva técnica del desafío “a priori”, y en la sección 5 se discuten sus propiedades y aplicaciones en otros sistemas de mensajería. Finalmente, en la sección 6 se concluye el artículo.

2. Problemas en los Sistemas de Mensajería

En los sistemas de mensajería, los principales problemas y al mismo tiempo causas de la aparición de *Spam* son la facilidad de acceso a las direcciones de los usuarios y la falta de autenticación de origen. Es muy sencillo obtener la dirección de un determinado usuario de forma automática, utilizando para ello “robots” (agentes pseudo-inteligentes) que o se encarguen de filtrar páginas web en busca de estas direcciones de contacto, o pregunten a servicios de localización de los propios servicios web. Una vez obtenidas las direcciones, en la mayoría de los casos se puede enviar un mensaje que no contenga información fiable sobre su procedencia real, ya que no existen mecanismos para poder autenticar al origen.

Todos estos problemas pueden encontrarse en los sistemas de correo electrónico, y más concretamente en su protocolo principal, SMTP. El protocolo SMTP fue introducido en 1982 [1], una época en la que mantener la seguridad de la red no suponía ningún problema ya que Internet estaba compuesta únicamente por miles de hosts. Actualmente el contexto es muy diferente, pero el protocolo sigue siendo el mismo (con ligeras modificaciones [2]).

En el protocolo SMTP, un mensaje consiste simplemente en una cadena de texto que contiene la siguiente información: origen, destino, servidores atravesados, mensaje, y cabeceras extras. El procedimiento para enviar de un mensaje de correo es sencillo: Un servidor de correo (cliente MTA) que maneja los correos del origen, contacta con el servidor de correo destino (servidor MTA) y le envía el mensaje. Es posible que un mensaje tenga que atravesar varios servidores MTA si el destinatario no es directamente accesible.

Sin embargo, un servidor MTA no puede averiguar quién le envió realmente el mensaje, debido a que un usuario malicioso puede tanto falsificar las cabeceras que indican quién fue el origen, como controlar o manipular un servidor de correo para que oculte quién fue el cliente MTA que envió el mensaje inicialmente. Como resultado, un Spammer (quien envía el *Spam*) puede enviar una cantidad ilimitada de *Spam* a cualquier usuario, y éste no podrá defenderse contra este ataque ya que no dispondrá de la información necesaria para poder bloquear el acceso de *Spam* a su cuenta de correo.

3. Trabajo Previo

El protocolo SMTP es un estándar que sirve como pilar a la infraestructura de correo electrónico de Internet. Como resultado, sería necesario planear una migración lenta y controlada (como está ocurriendo con IPv6) en caso de que el protocolo fuese cambiado. Por lo tanto, la mayoría de las investigaciones en el área de la lucha contra el *Spam* se centran en utilizar la información contenida en los mensajes (por ejemplo cabeceras) o en desarrollar aplicaciones que funcionen sin modificar el estándar.

Una cabecera capaz de proporcionar información útil es “Received:”. Esta cabecera ofrece una lista de los clientes y servidores MTA que han reenviado el mensaje a través de Internet, por lo que se puede comprobar si uno de esos servidores es una fuente de *Spam*. Existen algunos proyectos que tratan de clasificar este tipo de servidores [3,4]. No obstante, es posible bloquear clientes y/o servidores MTA inocentes.

Otra cabecera cuya información puede aprovecharse es la dirección del destinatario. Ésta dirección puede ampliarse con políticas de acceso o passwords. En sistemas basados en políticas de acceso [5], una política se codifica dentro de la dirección del destinatario, y si esta política no se cumple al llegar al servidor MTA, el mensaje de desecha. En los sistemas basados en passwords [6 – 8], la dirección del destinatario se amplía con una secuencia de caracteres que actúan como una password, la cual solo puede ser obtenida mediante una prueba de un gasto computacional [11]. Estas soluciones funcionan bien en algunos escenarios (por ejemplo cuando se utiliza una dirección de correo en entornos automatizados como foros de discusión), pero las direcciones de correo ampliadas son muy complejas, y son difíciles de recordar y utilizar para un ser humano.

Existen varios trabajos que se centran en analizar el contenido de un mensaje utilizando técnicas de inteligencia artificial (IA) y de análisis estadístico [9,10]. Como resultado de estos análisis se asigna una “puntuación” que distingue si un mensaje de correo proviene de un usuario legítimo o de un spammer. Sin embargo, estas técnicas pueden ocasionar falsos positivos (cuando un correos real es tratado como *Spam*) y falsos negativos (cuando un spammer modifica el formato de sus correos y éstos ya no se consideran *Spam*).

Otras soluciones existentes son las técnicas de micropago (micropayment), “desafío/respuesta” (Challenge/Response) y ofuscación. Los esquemas de micropago [11 – 14] evitan que los spammers envíen millones de correos basura, al ralentizar a los clientes MTA pidiéndoles calcular una función matemática compleja para poder comunicarse con el servidor MTA. Con todo, esta técnica es difícil de aplicar a dispositivos con recursos reducidos (como teléfonos móviles).

En las técnicas de “desafío/respuesta” [15, 16], siempre que un servidor MTA recibe un correo de un origen desconocido, éste responde automáticamente con un desafío. Una vez se responde al desafío, los correos procedentes de ese origen podrán alcanzar al destinatario. Sin embargo, estas técnicas introducen nuevos problemas (no solucionados hasta ahora), tales como la introducción de un tiempo de espera en la recepción de los mensajes o la posibilidad de ataques de denegación de servicio.

En los esquemas de ofuscación, las direcciones de correo se muestran al público de forma ofuscada (p. ej. nombre GUION apellido ARROBA servidor PUNTO dominio), y los usuarios que deseen utilizar esa dirección de correo deben “traducirla” primero. Un problema en esta solución es que las combinaciones para ofuscar una dirección de correo son limitadas, y una vez que la dirección esta capturada, el spammer puede enviar correos basura al destinatario sin mayor problema.

4. Propuesta de una nueva técnica: Desafíos “a priori”

4.1 Introducción

La técnica de desafíos “a priori” se basa en los mecanismos de “desafío/respuesta”, en el sentido de que ambas imponen al usuario que envía el mensaje un desafío que debe ser resuelto antes de acceder al correo del usuario destino. La particularidad de la técnica de desafíos “a priori” se encuentra en *cuando* se accede al desafío: Cuando un usuario desee enviar un correo, éste recogerá tanto la dirección de correo

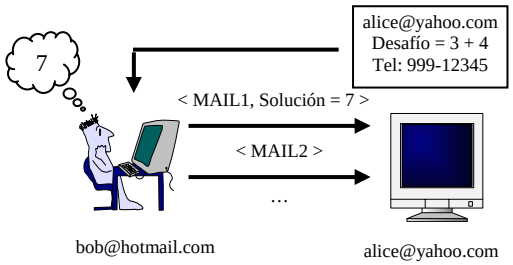


Fig. 1: Esquema básico del Desafío "a priori"

del destinatario como un desafío asociado a esa dirección de correo, al mismo tiempo. Una vez se resuelva el desafío, éste se incluirá dentro del mensaje (ver Fig. 1).

Cuando un correo que proviene de un remitente desconocido alcanza un buzón protegido, el servidor de correo comprobará que el mensaje incluye la solución al desafío. Si es así, el mensaje es admitido dentro del buzón, y el remitente se añade a una "*lista blanca*" (White-List) para que sus correos no necesiten incluir más el desafío resuelto (aún en el caso de que ese desafío sea cambiado).

El objetivo de esta técnica es el de comprobar que el remitente es realmente un ser humano. Esto es así debido a que los spammers utilizan programas automáticos tanto para obtener las direcciones de correo contenidas en páginas web o servidores de correo como para enviar el *Spam*. Sin embargo, es difícil para estos programas recoger un desafío asociado a una dirección de correo determinada, y les es aún más difícil tener el conocimiento semántico suficiente para resolver el desafío una vez obtenido. Por lo tanto, cuando un correo basura llegue a un buzón protegido, éste será descartado al no incluir la solución al desafío asociado a ese buzón.

En comparación con los esquemas de "desafío/respuesta", la técnica del desafío "a priori" conserva sus beneficios sin incluir sus defectos:

- En la técnica del "desafío/respuesta", existe un tiempo de espera para que el destinatario de un mensaje obtenga la solución a su desafío. Por otro lado, en el desafío "a priori", el desafío se encuentra disponible junto a la dirección del receptor, por lo que el remitente puede resolver el desafío y enviar su mensaje al receptor directamente.

- Si un spammer falsifica la dirección del remitente en sus correos, ese remitente recibirá los desafíos en caso de que el buzón del receptor esté protegido con la técnica del "desafío/respuesta", sufriendo un ataque de denegación de servicio [17]. Esto no ocurriría en el desafío "a priori", ya que los mensajes inválidos no provocan respuesta alguna.

- Un sistema de "desafío/respuesta" sólo puede funcionar con listas de correo si se incluyen algunas reglas específicas para cada lista de forma manual. En cambio, un sistema de desafío "a priori" puede manejar listas de correo y procesar mensajes de error sin ningún problema.

Otro beneficio del desafío "a priori" es la protección continuada que ofrece ante "robots" recolectores de direcciones. Una dirección de correo capturada por uno de estos programas no tiene utilidad a menos que se resuelva el desafío asociado a esa dirección. Y aún en el caso de que el desafío fuera resuelto, y la dirección fuese vendida (p. ej. en colecciones de CD), el dueño de la dirección puede modificar su desafío, haciendo que la combinación <email, solución> sea inútil.

La técnica del desafío "a priori" se integra de forma sencilla dentro de la infraestructura de correo actual, porque no obliga a cambiar ninguno de los protocolos de correo existentes (como POP3, IMAP, y SMTP). Puede incorporarse como un "plugin" dentro de cualquier servidor de correo, cuyas tareas serían las de proveer y mantener una serie de listas y reglas (secciones 4.3, 4.4 y 4.5) y las de interactuar con los dueños de los buzones de correo para tareas de mantenimiento (actualizar desafío y solución, modificar manualmente ciertas listas).

4.2 Obtención del Desafío

Cada cuenta de correo tiene un desafío asociado, y son los propietarios de esas cuentas quienes crean sus propios desafíos. Cada desafío puede ser actualizado en cualquier momento y tantas veces como su dueño desee. El grado de complejidad de los desafíos puede oscilar entre palabras o preguntas sencillas hasta sistemas complejos que solo un humano podría resolver [18].

En la mayoría de los casos un desafío se encuentra justo al lado de su dirección de correo asociada, de tal forma que cuando un posible remitente accede a la dirección de correo también puede recoger y resolver el desafío de forma inmediata. Sin embargo, en ciertos casos, puede que éste no se encuentre disponible de forma directa. En esos casos debe incluirse una URI que apunte a donde podría obtenerse ese desafío.

Ya que el desafío no se encuentra limitado a ofuscar una dirección de correo, la cual tiene una estructura fija (nombre, dominio), el usuario posee una mayor libertad en su creación. Cuando se almacena dentro de una página web, el desafío puede aprovechar el contenido que lo rodea (información personal, aspecto visual de la web). En entornos estáticos (p. ej. una tarjeta de visita) la solución al desafío puede incluirse directamente, ya que no hay peligro de que un spammer acceda a esa solución.

Finalmente, otra solución para obtener el desafío es la utilización de un servicio "majordomo" [21], donde un posible remitente pide a un servidor de correo cual es o donde está localizado el desafío de un usuario determinado. Eso sí, para prevenir que los spammers utilicen este servicio para recolectar direcciones de correo, el servicio debería devolver un desafío automáticamente generado para cada usuario no existente.

4.3 Estructuras de Datos

El desafío "a priori" requiere de ciertas estructuras de datos para poder funcionar. Las dos estructuras más importantes son el desafío "per se" (o una URI donde pudiera encontrarse) y la solución a ese desafío. Utilizando estas estructuras sería posible proporcionar el desafío actual a quienes lo necesiten y comprobar si la solución a un desafío es la correcta. Aparte, deben almacenarse las soluciones de antiguos desafíos.

Otras estructuras necesarias son la "*lista blanca*" (white-list), la "*lista de respuesta*" (reply-list) y la "*lista de alerta"* ("warning-list", específicamente diseñada para el desafío "a priori"). Cada una de esas estructuras contiene una lista de direcciones de correo y, adicionalmente, una fecha ("timestamp") para guardar el tiempo que una dirección de correo puede estar dentro de la lista.

"Lista Blanca". Los mensajes procedentes de remitentes incluidos en la *lista blanca* son inmediatamente admitidos dentro del buzón del usuario protegido, sin necesidad de comprobar la solución al desafío. Algunos remitentes pueden ser incluidos dentro de esta lista de forma manual si el usuario ya los conoce, evitando de esa forma que esos remitentes deban responder un desafío si el usuario ya confía en ellos.

"Lista de Respuesta". Esta lista contiene las direcciones de correo de aquellos usuarios a los que el propietario del buzón protegido ha enviado un correo, y aún no ha recibido respuesta. Es uso de esta lista se justifica de la siguiente forma: Si el propietario del buzón desea establecer una comunicación con otro usuario, sería innecesario requerirle una solución a un desafío.

"Lista de Alerta". La *lista de alerta* contiene las direcciones de correo de aquellos remitentes que han enviado un mensaje incluyendo una solución a un desafío antiguo. Debido a que es posible que un remitente solo haya tenido acceso a un desafío antiguo, la técnica del desafío "a priori" envía una respuesta automática a estos remitentes incluyendo el desafío actual. Cuando se incluye un remitente en esta lista, se indica que no debe recibir ninguna respuesta automática más en el futuro. En caso de cambiar el desafío actual, esta lista se vacía completamente.

4.4 Niveles de Seguridad

El desafío "a priori" puede ser configurado para trabajar en dos niveles de seguridad, *nivel alto* y *nivel bajo*. La diferencia entre ambos niveles de seguridad se encuentra en la forma de consultar la *lista de respuesta*.

El desafío "a priori" empieza trabajando en el nivel alto de seguridad. Este nivel implica que todas las consultas a la *lista de respuesta* se realizan buscando un par <usuario, dominio>, y que todas las coincidencias serán eliminadas. Por ejemplo, cuando se recibe un correo de bob@hotmail.com, se comprueban los campos "De:" y "Responder A:" del mensaje, y la *lista de respuesta* será consultada con el par <bob, hotmail.com>.

Por otro lado, en el nivel bajo de seguridad todas las consultas a la *lista de respuesta* se realizarán mediante el par <*, dominio>. De esta forma, si se recibe un correo de bob@hotmail.com, la *lista de respuesta* será consultada con el par <*, hotmail.com>.

La presencia de estos niveles de seguridad se debe a la existencia de cuentas de correo cuyas direcciones son distintas para enviar el correo y para recibir el correo. Esto suele ocurrir con las listas de correo, como se verá en la sección 5.1.

4.5 Funcionamiento del Desafío "a priori"

Ahora se procede a la explicación del funcionamiento de la arquitectura del desafío "a priori". Supondremos que existe un usuario B (remitente) que quiere enviar un mensaje al usuario A (destinatario), asumiendo (para simplificar la explicación) que el usuario A está utilizando el desafío "a priori" y que B no lo utiliza.

1. El servidor de A comprueba que la dirección de B se encuentre dentro de la *lista blanca*. En ese caso, el mensaje alcanza el buzón de A, y B recibe una confirmación si es el primer mensaje que envía a A.

2. En otro caso, si B se encuentra en la *lista de respuesta*, el correo alcanza el buzón de A y B es añadido a la *lista blanca*. Aquí hay que puntualizar que la consulta a la *lista de respuesta* es distinta dependiendo del nivel de seguridad al que esté funcionando el sistema (ver sección 4.4). En el caso de que el nivel de seguridad sea alto, la dirección de B se borra de la *lista de respuesta* (ya que A recibió la respuesta que esperaba de B).

3. En otro caso, se comprueba si el mensaje incluye la solución al desafío actual. Si es así, el correo alcanza el buzón de A, y B es añadido a la *lista blanca*. Adicionalmente, B recibe un correo de confirmación.

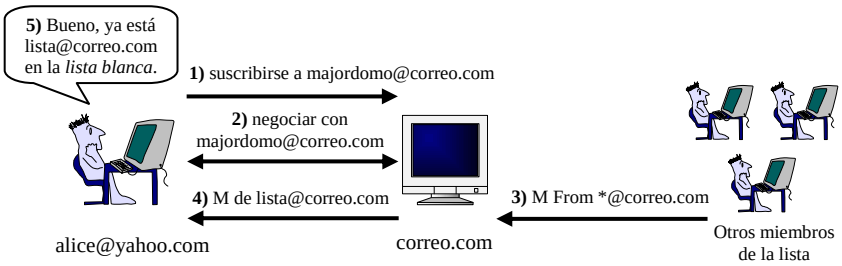


Fig. 2: Proceso de suscripción a una lista de correo.

4. En otro caso, si el mensaje incluye una solución a un desafío antiguo, se responde a B con un correo que incluya el desafío actual, pero solo si la dirección de B no se encuentra en la *lista de alerta*. Si se ha enviado la respuesta con el desafío actual, la dirección de B se añade a la *lista de alerta*.

5. En otro caso, el correo se desecha, sin enviar ningún tipo de respuesta a B. El problema de desechar un correo proveniente de un usuario legítimo se discute en la sección 5.3.

En este punto, cabe anotar que desechar un correo no significa que el usuario del buzón no pueda acceder a él. La técnica del desafío "a priori" puede configurarse para incluir una "puntuación" a los correos basura, de tal forma que el usuario pueda acceder a esos correos a través de su cliente de correo si así lo desea.

4.7 Escenarios

Existen dos posibles escenarios en el caso de que un spammer quiera enviar su correo basura a un usuario que esta protegido por la técnica del desafío "a priori".

Escenario 1. El spammer solo dispone de la dirección de correo del destinatario, pero no de su desafío. En ese caso, todo el correo basura enviado a un buzón protegido será automáticamente desechado, debido a la falta de una solución (tanto antigua como moderna) dentro del mensaje.

Escenario 2. El spammer solo dispone de la dirección de correo del destinatario, pero se identifica como un usuario existente dentro de la *lista blanca*, debido a los problemas de autenticación existentes en la infraestructura mundial de correo electrónico. Todas las técnicas que utilizan listas blancas comparten este inconveniente, pero no es un gran problema dado que un spammer debería encontrar una dirección de correo "valida" por cada una de las direcciones a las que quiere enviar *Spam*. Y para millones de direcciones, esto no es rentable.

Eso sí, podría parecer que un spammer solo necesita de un pequeño esfuerzo (resolver un desafío) para enviar todo el correo basura que desee a una dirección de correo en particular. También podría ocurrir que un grupo de spammers intercambiasen las soluciones de los desafíos que conocen para simplificar su tarea. Sin embargo, lo que los spammers persiguen es enviar millones de mensajes a millones de destinatarios. Y los desafíos son distintos por cada destinatario, y solo pueden ser resueltos por un ser humano. De esta forma, la tarea de espiar la red o utilizar mano de obra barata para obtener las soluciones a los desafíos no es rentable.

5. Discusiones

En esta sección se discuten como la técnica del desafío "a priori" funciona para usuarios de listas de correo, y también los problemas existentes en el acceso a un desafío. También se discute sobre como manejar adecuadamente mensajes de error, y de cómo aplicar nuestra técnica a sistemas como mensajería instantánea (IM) o blogs.

5.1 Manejo de Listas de Correo

Todas las listas de correo [19 – 21] poseen un mecanismo de registro similar: cuando un usuario desea registrarse dentro de una lista, ésta le envía un desafío para comprobar que quien ha enviado el mensaje es un ser humano. Este comportamiento hace imposible el manejo automático de listas de correo en sistemas de "desafío/respuesta".

Afortunadamente, existe una solución a este problema para la técnica del desafío "a priori", en la forma de los niveles de seguridad. Ya que todos los correos procedentes de una lista de correo pertenecen a un mismo dominio, es posible utilizar el nivel de seguridad bajo (ver sección 4.4) en el momento de empezar el registro dentro de la lista. De esta forma, todos los mensajes que se reciban durante el proceso de registro (desafíos incluidos) y que tengan una coincidencia en la *lista de respuesta* serán admitidos e incluidos dentro de la *lista blanca*. Finalmente, una vez que se reciba el primer correo de la lista, el usuario puede volver al nivel de seguridad alto (ver Fig. 2).

El riesgo de que un spammer entre en la *lista blanca* de un usuario mientras éste se encuentra en el nivel de seguridad bajo es pequeño, ya que la dirección de correo del spammer debe tener el mismo dominio que la de la lista de correo, y además un usuario se suele suscribir a muy pocas listas de correo al año.

Además, el usuario puede configurar el sistema para incluir las direcciones válidas en una *lista blanca* temporal cuando se funcione en el nivel de seguridad bajo, de tal forma que cuando el sistema pase al nivel de seguridad alto el usuario decida que direcciones de correo deben añadirse (manualmente) a la *lista blanca*.

5.2 Acceso al Desafío

Es evidente que existe un problema de disponibilidad si el desafío no se publica junto a su dirección de correo asociada. Si un usuario no puede obtener el desafío de otro usuario, sea porque acceder al desafío o al lugar que contiene el desafío no sea posible (p. ej. el remitente no puede acceder a Internet, o la página web que contiene el desafío esta bajo un ataque de denegación de servicio), es imposible que sus correos puedan alcanzar ese buzón protegido (sin ser marcados como *Spam*).

Debido a esa razón, es conveniente proporcionar tanto el desafío como una URI que apunte a donde ese desafío pueda obtenerse. De esta forma, si la URI no funciona, la solución al desafío, aunque éste no sea el actual, puede utilizarse para enviar un mensaje al destinatario (Si el desafío no es el actual el remitente recibirá un mensaje con el desafío que se está utilizando actualmente).

Finalmente, existe un problema de disponibilidad que es común tanto para los sistemas de desafíos "a priori" como para los sistemas de "desafío/respuesta". Un desafío que sea sencillo para un usuario concreto puede ser imposible de resolver para otro tipo de usuarios (por ejemplo, un usuario ciego no será capaz de resolver un desafío basado en imágenes).

5.3 Manejo de Falsos Positivos

Uno de los mayores problemas existentes en el desafío "a priori" ocurre cuando los correos de un usuario humano son desechados (sin enviar respuesta alguna) por el servidor de correo del destinatario protegido, al no incluir la solución a un desafío. Esto evita tanto que suba el tráfico en Internet como los ataques DoS causados por respuestas a los mensajes de spammers, pero a su vez un usuario que no sepa que un destinatario está protegido por un sistema de desafío "a priori" no será capaz de saber si sus mensajes han llegado a su destino o no.

Una posible solución consiste en definir un prefijo estándar para direcciones de correo protegidas por el mecanismo de desafíos "a priori". De esta forma, un remitente sabría que debe resolver un desafío para acceder al buzón del destinatario, y que si su primer mensaje es aceptado recibirá una confirmación.

Existe una solución alternativa en caso de que el desafío "a priori" se encuentre implementado en los servidores de correo. En esta solución, el usuario que envíe un mensaje no valido a un buzón protegido podrá recibir un mensaje de error gracias al protocolo de negociación de SMTP, sin que eso signifique un coste adicional para el servidor MTA que recibe el mensaje. Este protocolo funciona como sigue:

1. El cliente MTA del lado del remitente contacta con el servidor MTA del destinatario. Después de intercambiar mensajes de control, el servidor MTA permite al cliente MTA enviar el contenido del mensaje.

2. El cliente MTA envía el contenido del mensaje terminando con una simple ".". Después, el servidor MTA comprueba si el mensaje debe ser aceptado o rechazado. Si es rechazado, el cliente MTA recibe el mensaje "*554 Transaction failed* " (Transacción falllida).

3. Si la negociación fracasa, el cliente MTA genera un correo que incluya el mensaje original y el error enviado por el servidor MTA. Ese correo se envía al remitente original, en el caso de que este cliente MTA no maneje sus mensajes.

Cuando el servidor MTA comprueba si el correo es válido (paso 2), puede inspeccionar las cabeceras o contenidos del mensaje en busca de la solución al desafío del destinatario, ya que en este punto dispone de toda la información necesaria para realizar ese chequeo (origen, mensaje, destino). Si no hay solución al desafío, el servidor MTA puede devolver *"554 Transaction Failed: Solución al desafío errónea"* (indicando donde encontrar el desafío actual), y el cliente MTA generará un correo de error que incluirá automáticamente en el buzón del remitente.

5.4 Manejo de Mensajes de Error

Durante el curso del protocolo de negociación de SMTP, si un mensaje no puede llegar a su destinatario el cliente MTA debe enviar al remitente un correo que incluya las causas del error. Esos errores pueden ocurrir tanto por problemas de la cuenta destino (p. ej. cuota excedida) o por problemas administrativos o de seguridad (p. ej. solución de desafío no incluida).

Si el correo que avisa del error es generado por el cliente MTA que implementa el mecanismo de desafío "a priori", esto no supone ningún problema, ya que ese correo se incluye automáticamente en el buzón del usuario. Sin embargo, existe un problema en el caso de que el mensaje de error sea enviado al remitente original a través de un servidor MTA.

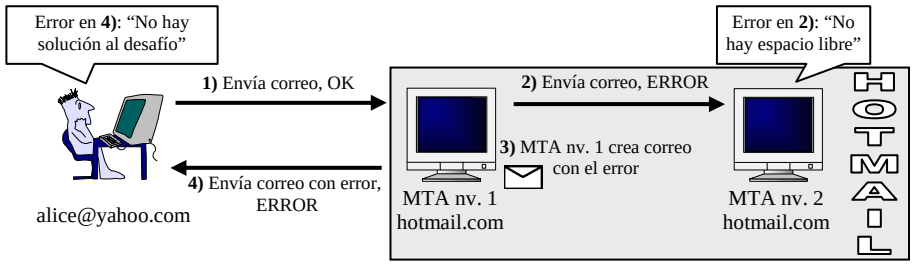


Fig. 3: Problemas en el manejo de mensajes de error

Un ejemplo de este problema puede verse en la Fig. 3. En el ejemplo, el error se produce en la MTA de nivel 2, por lo que la MTA de nivel 1 debe generar el mensaje de error. No obstante, el mensaje de error se envía a la MTA que maneja los mensajes del remitente original, y la MTA de nivel 1 no incluye la solución a ningún tipo de desafío – es, a fin de cuentas, una máquina.

Este problema puede resolverse gracias a dos principios: Primero, los mensajes de error pueden ser identificados gracias a la cabecera "message/delivery-status", e incluyen el mensaje original del remitente. Segundo, todos los correos tienen un número único que los identifica en la cabecera "message-ID".

De esta forma, cuando un mensaje de error llega a un buzón protegido, éste mensaje es aceptado si y solo si tanto la dirección del destinatario original como el número ID del mensaje original se encuentran en la *lista de respuesta*. Es por tanto necesario incluir el número ID de los mensajes dentro de la *lista de respuesta* si se desean manejar correctamente los mensajes de error. Dado que para obtener este número ID sería necesario interceptar un correo en su viaje hacia el servidor MTA, los spammers no pueden sacar provecho de este mecanismo.

Hay que hacer notar que un spammer podría realizar un ataque DoS a un buzón no protegido si enviara mensajes a un servidor con una dirección de origen falsificada, en el escenario expuesto en la Fig. 3. Esta situación se evita incorporando los mecanismos del desafío "a priori" en las MTAs de nivel 1.

5.5 Protección de Otros Sistemas de Mensajería

5.5.1 *Spam* de Mensajería Instantánea (IM)

Los sistemas de mensajería instantánea (IM) proporcionan servicios de comunicación simple (texto) o compleja (audio/video) entre dos extremos, y servicios de localización entre un grupo de usuarios denominados "lista de amigos".

Un usuario debe registrarse primero dentro de un servicio de mensajería instantánea para poder contactar con otros usuarios. Además, los usuarios tienen mecanismos que les permiten comprobar quién quiere comunicarse con ellos, y pueden prohibir el acceso a usuarios sospechosos. Por esa razón, el *Spam* no es un problema común en estos sistemas.

No obstante, existen ciertos servicios de IM que sufren el problema del *Spam*, como el servicio World-Wide Pager de ICQ [22]. Estos servicios permiten que usuarios anónimos envíen un mensaje instantáneo, utilizando un formulario HTML, a cualquier usuario. Como hay autenticación de origen, existen programas automáticos que permiten enviar *Spam* a usuarios de IM en tiempo real.

Dado que estos servicios de IM están incluidos en páginas web, la técnica del desafío "a priori" puede ser utilizada, permitiendo a los usuarios ofrecer un desafío a aquellos que quieran enviarles un mensaje. De esta forma, el *Spam* de mensajería instantánea no sería rentable, tal y como se ha explicado en este artículo.

5.5.2 Blog *Spam*

Los Weblogs (o simplemente Blogs) son un tipo de aplicación web en el que uno o más usuarios escriben información (no modificable) que más tarde podrá ser accedida por otros usuarios. Una de las características más interesantes de los blogs es que permite que los visitantes escriban comentarios sobre la información incluida en cualquier parte del blog.

Sin embargo, es posible que un blog reciba *Spam*, en forma de un comentario corto que incluye un enlace a una página web, la cual suele anunciar un producto fraudulento. El objetivo de este tipo de *Spam* es el de aumentar la importancia de esas páginas web en buscadores como google, y provocan que los usuarios legítimos tengan dificultades en leer comentarios que merezcan la pena.

Una solución desarrollada por google [23] consiste en incorporar automáticamente a la etiqueta HREF de HTML la opción NOFOLLOW, de tal forma que los enlaces existentes dentro de un comentario no servirán a la hora de contar la prioridad de la página web enlazada. Sin embargo, es posible que esto no acabe con el *Spam* debido a la existencia de Blogs sin proteger y a los bajos conocimientos técnicos de los spammers.

La técnica del desafío "a priori" puede utilizarse también para proteger a los blogs del *Spam*. Si fuera

necesario responder a un desafío antes de poder enviar un comentario, los programas automáticos de envío de *Spam* dejarían de funcionar (como se ha discutido durante todo este articulo), y los comentarios se verían libres de *Spam*.

6. Conclusión

En este artículo, se ha presentado una técnica denominada desafío "a priori" para controlar el *Spam* del correo electrónico, basada en los mecanismos de "desafío/respuesta" pero sin ninguno de sus problemas, y capaz de proteger otros sistemas de mensajería como la Mensajería Instantánea y los Blogs.

Esta técnica también puede ser utilizada conjuntamente con otras soluciones contra el *Spam*. Así se deja la puerta abierta a otras soluciones como las de análisis de contenido. Al mismo tiempo, es posible integrar nuestra técnica con sistemas de autenticación de origen como DomainKeys [24] o IBE [25], evitando los problemas de autenticación que surgen en el manejo de la *lista blanca*.

Referencias

[1] J. Postel. *Simple Mail Transfer Protocol*. RFC 821, Internet Engineering Task Force, Agosto 1982.

[2] J. Klensin. *Simple Mail Transfer Protocol*. RFC 2821, Internet Engineering Task Force, Abril 2001.

[3] RBL. http://mail-abuse.org/rbl/.

[4] SBL. http://spamhaus.org/.

[5] J. Ioannidis. *Fighting Spam by Encapsulating Policy in Email Addresses*. In Proceedings of NDSS'03 (Network and Distributed System Security), Febrero 2003.

[6] E. Gabber, M. Jakobsson, Y. Matias, and A. Mayer. *Curbing Junk E-Mail via Secure Classification*. In Proceedings of FC'98 (Financial Cryptography), pages 198--213, Febrero 1998.

[7] R. J. Hall. *How to Avoid Unwanted Email*. Communications of the ACM, 41(3):88-95, Marzo 1998.

[8] L. F. Cranor and B. A. LaMacchia. *Spam!*. Communications of the ACM, 41(8):74--83, Agosto 1998.

[9] M. Sahami, S. Dumais, D. Heckerman, and E. Horvitz. *A Bayesian Approach to Filtering Junk Email*. In Proceedings of AAAI'98 Workshop on Learning for Text Categorization, Julio 1998.

[10] P. Cunningham, N. Nowlan, S. J. Delany, and M. Haahr. *A Case-Based Approach to Spam Filtering that Can Track Concept Drift*. In Proceedings of ICCBR'03 Workshop on Long-Lived CBR Systems, Junio 2003.

[11] C. Dwork and M. Naor. *Pricing via Processing or Combatting Junk Mail*. In Proceedings of Crypto'92, pages 139--147, Agosto 1992.

[12] C. Dwork, A. Goldberg, and M. Naor. *On Memory-Bound Functions for Fighting Spam*. In Proceedings of Crypto'03, pages 426--444, Agosto 2003.

[13] M. Abadi, A. Birrell, M. Burrows, F. Dabek, and T. Wobber. *Bankable Postage for Network Services*. Proceedings of the 8th Asian Computing Science Conference, Mumbai, India, Diciembre 2003.

[14] Penny Black Project, Microsoft Research. http://research.microsoft.com/research/sv/PennyBlack/.

[15] SpamArrest. http://spamarrest.com/faq/.

[16] SpamCap http://www.toyz.org/cgi-bin/wiki.cgi?SpamCap.

[17] J. Mirkovic, J. Martin, and P. Reiher. *A Taxonomy of DDoS Attacks and DDoS Defense Mechanisms*. University of California, Computer Science Department, Technical Report #020018.

[18] L. von Ahn, M. Blum, N. J. Hopper, and J. Langford. *CAPTCHA: Using Hard AI Problems for Security*. In Proceedings of Eurocrypt'03, pages 294--311, Mayo 2003.

[19] Ezmlm Mailing List. http://www.ezmlm.org/.

[20] Mailman Mailing List. http://www.list.org/.

[21] Majordomo Mailing List. http://www.greatcircle.com/majordomo/.

[22] ICQ Pager. http://www.icq.com/panels/messagepanel/.

[23] Google Blog. *Preventing Blog Spam (Enero 18,2005)*. http://www.google.com/googleblog/2005/01/preventing-comment-spam.html

[24] Yahoo DomainKeys. http://antispam.yahoo.com/domainkeys/.

[25] D. Boneh and M. Franklin. *Identity Based Encryption from the Weil Pairing*. Crypto'01, pages 213-229, Agosto 2001.

Monitorización activa de altas prestaciones mediante la plataforma paneuropa ETOMIC

E. Magaña, U. Alonso, F. Astiz, D. Morató, M. Izal, F. Naranjo y J. Aracil
Departamento de Automática y Computación
Universidad Pública de Navarra
C/ Campus Arrosadia, 31006 Pamplona
Teléfono: 948 169853 Fax: 948 168924
E-mail: eduardo.magana@unavarra.es

Abstract *In this paper we present the first set of active measurements that we have made using the ETOMIC system. ETOMIC is a paneuropean traffic measurement infrastructure with GPS-synchronized monitoring nodes. Specific hardware is used in order to provide high-precision transmission and reception capabilities. Besides, the system is open and any experiment can be executed. Internet measurements with high infrastructure requirements are now possible like one-way delay, routes and topology changing, congestion detection and virtual path aggregation detection. We will explain the results and how easy is to implement these measurements using the tools provided by ETOMIC, specially the API for using the specific sending and receiving capabilities.*

1 Introducción

Uno de los grandes campos de investigación en la actualidad es la medición de red mediante técnicas de monitorización activa o pasiva, de forma que se pueda extraer suficiente información como para poder realizar un dimensionamiento correcto del crecimiento de las troncales, modelar el tráfico de los usuarios o detectar las necesidades de calidad de servicio de nuevas aplicaciones. Sin embargo, conforme aumenta la velocidad de las troncales y redes de acceso se hace necesario disponer de herramientas de monitorización más sofisticadas que las que habitualmente se han venido utilizando.

De esta manera, si nos queremos enfrentar al problema de realizar medidas en Internet nos encontramos con los siguientes requisitos. En primer lugar será necesario una alta precisión en las marcas temporales de las medidas a realizar, sobre todo conforme se trabaje con mayores anchos de banda. Además, si se quiere realizar mediciones en una Internet global, necesitaremos mecanimos que provean la sincronización precisa entre equipos dispersos geográficamente, de manera que se pueda realizar una interpretación correcta de los resultados. Por otro lado, para lo que a monitorización activa se refiere [1], se hará necesario disponer de hardware específico que permita generar ráfagas de paquetes con alta precisión temporal en el espaciado entre paquetes o que permita la recepción de paquetes sin pérdidas. Estos requisitos quedarán por lo general fuera de las capacidades de plataformas PC de bajo coste habitualmente utilizadas.

Precisamente el proyecto integrado EVERGROW[1] del VI Programa Marco de la Unión Europea tiene como objetivo fundamental el crear unas bases de conocimiento de la evolución de Internet hasta el año 2025. Aplicando teorías de sistemas complejos se hace necesario disponer de datos reales de la Internet actual. Para obtener esos datos, dentro de este proyecto integrado se enmarca la plataforma de monitorización ETOMIC (http://www.etomic.org, *European Traffic Observatory Measurement InfrastruCture*). El sistema ETOMIC consiste en un sistema central de gestión (CMS, Central Management System) que gestiona y supervisa nodos de monitorización distribuídos por toda Europa. En la actualidad se dispone de 12 nodos y se pretende llegar a tener 50 en localizaciones escogidas de toda Europa, principalmente en universidades, centros de investigación, operadoras y empresas de telecomunicaciones.

Las características principales de ETOMIC [2, 3] son las siquientes:

- Ofrece una plataforma de monitorización abierta a la comunidad investigadora, de manera que cualquier investigador pueda obtener una cuenta en el sistema y acceder a todos los recursos.
- Provee monitorización activa y pasiva, si bien en la fase inicial está enfocado a monitorización activa.
- Es totalmente reconfigurable, dejando al investigador la posibilidad de realizar cualquier tipo de medida al permitirle correr en los nodos el software que desee. Se dispone además de una selección básica de programas que permiten realizar las medidas más habituales.
- Todo ello dentro de un esquema hardware de alta precisión, tanto en temporización como en sincronización, para recepción y emisión de paquetes sobre la red.

[1]Este trabajo ha sido financiado por el Proyecto Integrado Evergrow (contrato 001935) del Programa FP6/IST/FET de la Comisión Europea. http://www.evergrow.org

Para realizar el presente estudio se ha utilizado la plataforma ETOMIC mostrando algunas de las medidas que es posible realizar y con ello la potencia real del sistema.

2 Plataforma ETOMIC

La plataforma ETOMIC provee a los investigadores de un interfaz web de acceso al sistema. Este interfaz es servido por el CMS que es el encargado de gestionar la realización de los experimentos que se deseen en los nodos distribuidos por Europa.

Cada uno de los nodos está dotado del siguiente hardware que los convierte en herramientas de alta precisión:

- Plataforma PC: un PC con Debian GNU/Linux proveerá capacidades de almacenamiento de datos en disco, comunicación con el CMS, watchdog y el entorno para la ejecución de experimentos. Dispone de una tarjeta Ethernet para comunicación de gestión con el CMS, pero también utilizable para los experimentos.

- GPS Garmin 35 HVS: permite dotar al sistema de alta precisión en las marcas temporales de paquetes recibidos y en el espaciado de paquetes en emisión, así como sincronización de los relojes de todos los nodos. La precisión del sistema será uno de los aspectos estudiados en este trabajo.

- Tarjetas Endace DAG 3.6GE [4]: tarjetas Ethernet 10/100/1000 con conexión al GPS, permite colocar un *timestamp* por hardware a la recepción de los paquetes y definir la temporización de ráfagas de paquetes en emisión también por hardware, consiguiendo cotas de precisión muy por encima de los sistemas convencionales. Además se basan en una arquitectura de memoria compartida con lo que eliminan la sobrecarga de interrupciones de los interfaces de red habituales.

- Convertidor serie: permite adaptar las señales del GPS (RS-232) con la entrada de la tarjeta Endace DAG 3.6GE (RS-422) y el puerto serie del ordenador.

En la Fig.1 se muestran los componentes específicos del sistema: tarjeta Endace DAG, GPS y conversor. Para el proyecto se ha desarrollado un firmware para las tarjetas Endace DAG 3.6GE que ofrece funcionalidades de generación de tráfico (el firmware convencional únicamente tiene capacidades en recepción), proveyendo un esquema para la generación de ráfagas de hasta 256 paquetes donde es posible definir el tamaño y espaciado entre paquetes con resolución del orden de nanosegundos.

Para facilitar la labor del investigador se ha realizado un API que permite aprovechar las funcionalidades de las tarjetas Endace DAG a través de un interfaz muy sencillo para generación y recepción de

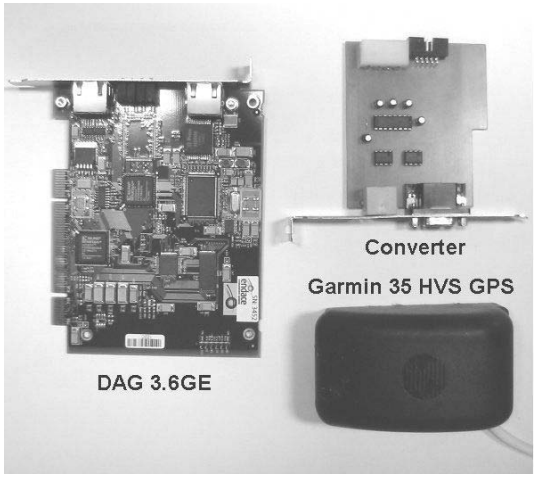


Figura 1: Endace DAG, GPS y conversor

paquetes, tanto para modo activo como pasivo. El API permite comprobar los programas desarrollados para las tarjetas Endace DAG sin disponer de ellas, realizando una emulación por medio de una tarjeta de red Ethernet convencional en el PC del investigador. Este API será de utilidad para aquellos investigadores que deseen lanzar sus propios programas en cada nodo, siempre dentro de las limitaciones de un usuario normal sin permisos de administración.

Actualmente se encuentran desplegados nodos en Suecia, Inglaterra, Alemania, Francia, Hungría, Italia, Israel y España. El CMS se encuentra localizado en la Universidad Pública de Navarra, Pamplona. Estos nodos requieren de una instalación especial debido a la necesidad de colocar el GPS con buena visibilidad aérea, en la mayoria de los casos conseguida colocándolo en el tejado (normalmente a decenas de metros de distancia del nodo). En la Fig. 2 se muestra una captura de la web de información del sistema con la disposición de los nodos operativos. Este interfaz web provee a los usuarios de todo lo necesario para realizar sus experimentos, desde subir y compilar los programas hasta reservar los nodos y seleccionar los programas a correr en cada nodo. Las fases típicas de definición y ejecución de los experimentos son las siguientes:

- Subir al CMS ficheros ejecutables, código fuente o ficheros de datos. Una vez en el CMS tendremos accesibles esos ficheros para programar los experimentos en los nodos.

- Crear un *bundle*. El bundle define las características de nuestro experimento sin concretar el momento en el que se va a realizar. Para ello deberemos seleccionar qué nodos necesitamos para el experimento y definir los ficheros que es necesario subir a cada uno de ellos, así como los puntos de ejecución que queremos insertar para lanzar nuestros programas con la temporización que consideremos oportuna.

- Definir un experimento. Consiste en asignar

la franja temporal en la que queremos lanzar determinado bundle definido con anterioridad.

- El CMS se encarga de esperar y planificar la ejecución del experimento de manera automática sin intervención del investigador. Antes de la hora programada para el experimento se encargará de subir los ficheros necesarios para cada nodo, y en el momento de cada ejecución lanzar los programas adecuados. Cuando termine el tiempo para el experimento se bajará los resultados de cada nodo al CMS.

- Descarga de resultados. El investigador, cuando haya finalizado el experimento y lo estime oportuno, podrá utilizar el interfaz web para bajarse los resultados de su experimento desde el CMS.

Toda la interacción del investigador y la plataforma ETOMIC se realiza a través del interfaz web del CMS, facilitando en gran medida la realización de experimentos complejos en múltiples nodos simultáneamente.

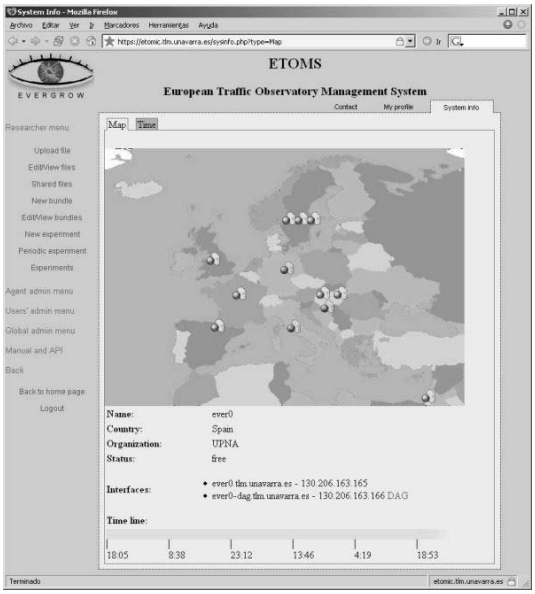


Figura 2: Interfaz web para el investigador

3 Precisión del sistema ETOMIC

En las redes actuales, un requisito previo para obtener mediciones correctas es disponer de un instrumento de medida con una resolución suficiente para manejar paquetes que se transmiten a altas velocidades. Más concretamente, la mayor parte de los enlaces troncales de las redes Rediris y Geant están formados por enlaces de 2.5 y 10 Gbps. Además de una buena resolución se requiere que los nodos estén sincronizados, o dicho de otra forma, los nodos tienen que compartir la misma hora con una gran precisión.

Como ejemplo de la resolución necesaria, un paquete Ethernet de tamaño máximo en un enlace Gigabit tarda en transmitirse 12 microsegundos y en general se requiere una resolución de decenas de nanosegundos. Es importante disponer de gran resolución cuando, por ejemplo, se quieren detectar espaciados entre paquetes enviados de forma inmediatamente consecutiva (*back-to-back*). El envío de paquetes *back-to-back* se puede utilizar para la estimación del ancho de banda.

La sincronización entre los nodos se resuelve gracias al sistema GPS, el cual además de proveer los conocidos servicios de localización globales también ofrece servicio de sincronización. La sincronización es necesaria para medir correctamente el tiempo que tarda un paquete desde que se manda desde un nodo origen hasta que llega al nodo destino (retardo). El retardo de los paquetes además de indicar la velocidad de transmisión de los paquetes ofrece información sobre el nivel de congestión de las redes. Aunque la magnitud en las medidas de retardo es del orden de decenas de milisegundos, la sincronización debe ofrecer una mayor precisión o en caso contrario medidas como el espaciado entre paquetes *back-to-back* obtendrían valores oscilantes no tolerables.

El procedimiento básico para realizar mediciones es el de establecer una marca de tiempo cuando se transmite y otra cuando se recibe. Estas marcas de tiempo se deben establecer con el mínimo retardo y variabilidad posible. Para ello no se puede confiar en hardware convencional o en sistemas operativos de propósito general. El hardware convencional no ofrece servicios de marcas de tiempo en los paquetes y tienden a ofrecer unos tiempos de respuesta variables. Los sistemas operativos de propósito general no están diseñados para ofrecer marcas de tiempos con una alta precisión y en cualquier caso están supeditados a la no predecibilidad temporal del hardware convencional. Los errores de *timestamp* de este tipo de plataformas están alrededor de 0.1ms [5] con respecto al reloj real de la máquina, sin contar las desviaciones de éste y la falta de sincronización entre diferentes equipos.

La solución utilizada para resolver todos estos problemas ha sido utilizar hardware de propósito específico, concretamente se han utilizado tarjetas Gigabit-Ethernet DAG 3.6GE de Endance. Las principales características de estas tarjetas de red son:

- Utilizan GPS para su sincronización.

- Establecen marcas de tiempo en los paquetes en emisión y en recepción con una resolución aproximada de 60 ns (exactamente $1/2^{24}$ segundos).

- Permite la programación del envío de paquetes con un espaciado concreto con la resolución mencionada en el punto anterior de forma que el sistema operativo no influye en la temporización de los paquetes.

La evaluación de la precisión se ha hecho realizando medidas del tiempo de retardo entre dos nodos a través de una red con retardo constante. Debido a la imposibilidad de conectar dos tarjetas DAG con un

cable cruzado se ha utilizado un concentrador y se ha comprobado que el retardo a través del concentrador es suficientemente estable. Más concretamente, en la evaluación se comparará una configuración en la que las tarjetas se sincronizan entre sí, es decir sin GPS y compartiendo la misma fuente de sincronización (una de las propias tarjetas en modo *master*) y otra configuración en la que cada tarjeta está sincronizada con su propio GPS. El hardware utilizado ha sido: tarjetas de red DAG 3.6GE de Endance, concentrador Fast-Ethernet OfficeConnect DualSpeed Hub 8 (3C16753) de 3Com, receptores GPS Garmin 35-HVS y tarjetas conversoras serie.

La primera prueba consiste en validar la prueba diseñada, esto se realiza utilizando la configuración sin GPS y observando por una parte la variación del tiempo de retardo y por otra la diferencia de tiempo entre paquetes en tiempo de emisión y recepción. Ambos tiempos deben de ser el mismo con una alta precisión, los resultados obtenidos en esta prueba indicarán la precisión de los resultados que se obtendrán en pruebas posteriores.

En las pruebas se obtuvo que el retardo promedio fué de 1276 ns, que la diferencia entre el retardo máximo y mínimo fué de 299 ns (menos de 30 bits en Fast-Ethernet.) y la desviación estándar 41 ns. Nos referiremos al valor del retardo promedio obtenido en esta prueba como *retardo calibrado* y será utilizado cuando comparemos los valores de retardo obtenidos cuando se realicen medidas utilizando GPS.

Por otra parte, la diferencia de tiempo entre paquetes en tiempo de emisión y recepción es como mucho dos veces la resolución del reloj de las tarjetas, es decir ± 120 ns, obteniendo el mismo tiempo tanto en emisión como en recepción el 81 % de las veces.

Una vez validada la prueba diseñada, se procede a evaluar la bondad de la sincronización mediante GPS. Cabe destacar que el dispositivo GPS utilizado es de bajo coste y no está diseñado específicamente para obtener un rendimiento óptimo en cuanto a sincronización de tiempos.

Se realizaron pruebas de 4 horas en las que se enviaban 2 paquetes por segundo, o lo que es lo mismo 28800 muestras. Una vez obtenidos los valores de retardo de los paquetes, se les resta el valor del *retardo calibrado* para obtener una medida del error. En la Fig. 3 se pueden apreciar los resultados obtenidos, en los que se obtiene un error promedio es de 750 ns con una desviación estándar de 140 ns.

En general es esperable que el desplazamiento entre los relojes sea del orden de decenas o centenas de nanosegundos. Algunas razones para ello son:

- Las prestaciones del dispositivo GPS.
- Diferentes longitudes de los cables de conexión con el GPS. No es extraño que pueda haber diferencias en cables de 100 metros que equivalgan a 500 ns en tiempo de propagación de la señal eléctrica.

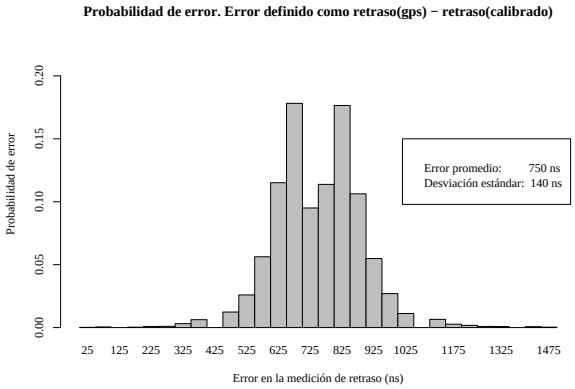


Figura 3: Histograma de la probabilidad de error utilizando GPS.

- La temperatura afecta a los relojes de cuarzo de los dispositivos electrónicos. En este caso concreto, los dispositivos GPS se encuentran en el exterior, afectándoles el clima y las tarjetas DAG pueden estar en entornos de ordenadores donde no suele haber buena ventilación. Es decir los relojes que proporcionan las marcas de tiempo de los diferentes nodos pueden tener unas derivas diferentes según la temperatura ambiente del reloj.
- El clima exterior (nieve, lluvia/humedad, nubes y temperatura) también afecta a la propagación de la señal de los satélites.

Debido a las cotas de precisión de los relojes, tanto la magnitud del desplazamiento como la de la oscilación del error no es problemático para medir el tiempo de retardo entre los nodos, ya que como se mencionó previamente tienen un retardo que está en el orden de milisegundos.

Un desplazamiento constante del error no afecta a las medidas de espaciados entre paquetes, sin embargo la oscilación del error sí. Esto último se debe a que el espaciado de los paquetes puede ser inferior a los microsegundos. A este respecto los resultados anteriores se deben interpretar como “cada 500 ms hay una oscilación con una desviación típica de 140 ns”. Es decir, en espaciados inferiores a estos 500 ms el error será proporcionalmente menor hasta llegar a la resolución del reloj (aprox 60 ns). En el caso de utilizar trenes de paquetes enviados *back-to-back* con una longitud mayor se deberá utilizar algún procedimiento estadístico para minimizar el error.

4 Medición del retardo *one-way* y pérdidas

4.1 Función de densidad

Algunas aplicaciones (Voz sobre IP, sesiones interactivas, chat, etc.) son sensibles al retardo en un sentido y a sus variaciones, así como a las

perdidas [6] [7]. ETOMIC permite realizar varios tipos de medidas del retardo, sin necesidad de tecnicas de estimación [8]. Gracias a la sincronización de los agentes basta con leer en recepción la marca de tiempo introducida en cada paquete en transmisión. Estas marcas temporales las introduce por hardware la propia tarjeta Endace DAG 3.6GE.

Para realizar este tipo de medidas se han desarrollado dos programas. Uno de ellos (trafgen) manda paquetes con las características que se deseen (protocolo, puerto, ttl, etc.) al destino que se indique. Los tiempos entre 2 paquetes consecutivos se definen mediante un fichero, o indicando al programa que sigan una distribución determinada. El tamaño de los paquetes puede indicarse de la misma manera. El segundo programa (sink) filtra los paquetes que se reciben en la tarjeta DAG siguiendo las reglas que se deseen. De los paquetes que han pasado el filtro se obtiene la información necesaria.

Configuramos trafgen de forma que envíe con tiempo entre paquetes constante. El tamaño de los paquetes es constante e igual a 46 bytes. Utilizamos el protocolo UDP y los puertos de origen y destino son ambos 80.

El programa sink filtra los paquetes que no sean UDP con puertos de origen y destino 80. Entre toda la información que nos ofrece se encuentra el retardo en un sentido, calculado gracias al *timestamp* origen incluído en el paquete.

La Fig. 4 muestra la densidad de probabilidad de los retardos medidos desde un nodo en Pamplona (España) hasta otro situado en Birmingham (Gran Bretaña). Este experimento en concreto se ha realizado con un ratio constante de 150 paquetes por segundo. El tamaño de los paquetes es 46 bytes. En esta medida además de apreciar la distribución que sigue la medida del retardo se puede comprobar la precisión del sistema: nótese la escala del eje de abscisas, con cientos de puntos de medida por cada intervalo de 200ns.

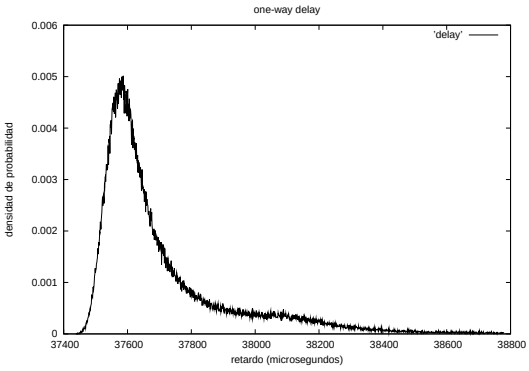


Figura 4: Retardo en un sentido Pamplona - Birmingham

Cabe destacar que el retardo no es el mismo en ambos sentidos. La Fig. 5 muestra la función densidad de probabilidad del retardo Pamplona-Birmingham en ambos sentidos, comprobándose la gran diferencia de la media en ambas, debido posiblemente a diferente congestión en los enlaces en cada sentido.

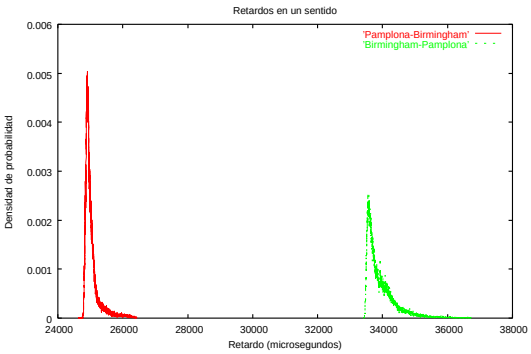


Figura 5: Retardos en ambos sentidos Pamplona - Birmingham

4.2 Detección de congestión

La medida de las perdidas en un enlace puede utilizarse para detectar congestión o cuellos de botella. Para realizar esa tarea se envían dos rafagas de paquetes utilizando las mismas aplicaciones explicadas anteriormente. Los paquetes se envían desde un nodo en Budapest (Hungría) hasta un nodo en Jerusalem (Israel).

Transmitiendo ráfagas de paquetes a velocidades de hasta 8 Mbit/s no hay prácticamente perdidas de paquetes y el retardo en un sentido tiene un valor constante para todos los paquetes. En la Fig. 6 se envían 2 rafagas a una velocidad de 12 Mbit/s (representadas por la línea continua inferior en la figura). En este caso se ve claramente cómo un elemento está encolando los paquetes y haciendo que aumente su retardo. Cuando el buffer de ese dispositivo de red se llena comienza a descartar paquetes sistemáticamente (aspas sobre el valor 20 de retardo de la Fig. 6 indican paquetes perdidos) hasta que se vacía parcialmente y vuelve a cursarlos. Es decir, el mecanismo de planificación de la cola aplica cierta técnica de histéresis. En las dos ráfagas el comportamiento es idéntico y repetible en diferentes experimentos. Si además nos fijamos en la pendiente del retardo se observa cómo el cuello de botella está marcado por un tramo que está funcionando a 10 Mbit/s en el camino.

Este tipo de medidas, con ráfagas de paquetes, además de permitir la detección de congestión o cuellos de botella, sirven para obtener ciertas propiedades del camino [9]. En este caso se observa un comportamiento de un router que no corresponde con los algoritmos más habituales de descarte de paquetes [10]: FIFO, RED, CBQ, etc.

4.3 Efectos multicamino y del hardware de red

El protocolo IP utilizado por todos los dispositivos en Internet es un protocolo conceptualmente sencillo en el que cada paquete se reenvía de forma independiente

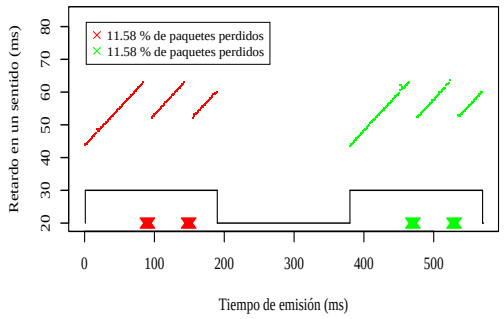


Figura 6: Ráfagas a 12 Mbps

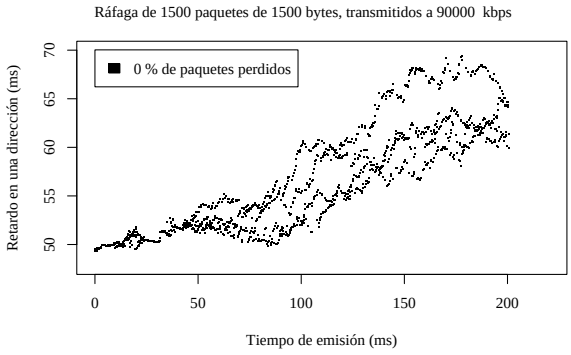


Figura 7: Incremento del retardo de los paquetes en 4 tendencias, produciendo reordenación.

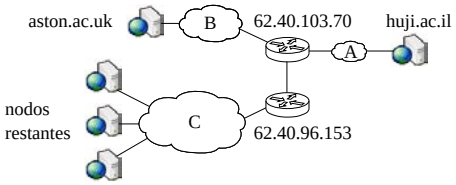


Figura 8: Topología de red de los nodos involucrados en el diagnóstico de la reordenación de paquetes hacia el nodo en Israel (huji.ac.il).

entre los nodos intermedios hasta alcanzar el destino. IP no requiere que los paquetes lleguen de forma orientada, ni siquiera tiene cuidado de que lleguen, se dice que IP es un protocolo que no garantiza el servicio (*best effort*). TCP, encapsulado por encima de IP, es el protocolo de transporte mas utilizado en Internet (más del 90 % [11]). Este protocolo es utilizado en una transmisión únicamente por el emisor y el receptor (no los nodos intermedios) y está diseñado para que el receptor avise al emisor de cuándo se producen pérdidas de paquetes. En este caso el emisor reenvía los paquetes perdidos y entiende que las pérdidas se producen por congestión de la red, por lo que baja el ritmo de emisión. Cuando se produce reordenación, aunque no se produzcan pérdidas TCP puede interpretar la reordenación como pérdidas: la conexión se vuelve más lenta porque utiliza la red de forma ineficiente (al reenviar paquetes innecesarios) y la tasa de transferencia se reduce ya que el emisor interpreta que hay congestión en la red.

Durante la realización de experimentos se comprobó que en el nodo de la universidad Hebrea de Jerusalén (huji.ac.il) solía producirse reordenación. Se puede apreciar este efecto en la Fig. 7, en la cual paquetes enviados de forma consecutiva obtienen un retardo muy diferente unos respecto a otros. Si nos fijamos en esta figura se puede comprobar cómo parece haber 4 líneas de tendencias diferentes en los retardos.

Se realizaron pruebas enviando tráfico desde todos los nodos hacia huji.ac.il. Se comprobó que el efecto de reordenación mostrado en la Fig. 7, se producía con mayor o menor intensidad desde todos los nodos excepto el de la Universidad de Aston (aston.ac.uk).

Utilizando la información de la topología de red se comprobó que el efecto no se producía por un cambio continuo en las rutas (*route flapping*, habitualmente debido a una mala configuración de los protocolos de encaminamiento) o porque se estuviera haciendo un reparto de carga de tráfico entre los diferentes caminos hacia huji.ac.il.

A partir de la misma información topológica se comprobó que todos los nodos, cuando envían tráfico hacia huji.ac.il, comparten el camino a partir del nodo 62.40.103.70 (nube C en la Fig. 8). Dado que desde aston.ac.uk no se produce reordenamiento, la causa del fenómeno debe estar en un punto intermedio compartido entre el resto de nodos y no por aston.ac.uk (nube C). El causante puede ser cualquier salto intermedio compartido por el resto de los nodos y que esté situado entre 62.40.103.70 y éstos situados en la nube B. Afortunadamente, únicamente 62.40.96.153 (situado inmediatamente a continuación) cumplía esta condición.

Descartadas las causas de cambio de ruta, quedan las siguientes posibilidades:

- El reordenamiento de paquetes en 4 tendencias diferentes puede estar causado por el mal reparto de paquetes en un enlace lógico agregando 4 enlaces físicos (por ejemplo un OC-12 se puede componer con 4 enlaces OC-3).
- Una mala planificación de los paquetes en ese nodo.

El primer punto no es: observando el mapa de capacidades de Geant [12] el enlace es inferior a un OC-3 y este constaría de 3 (y no 4) enlaces físicos OC-1. Contactos con Dante (entidad que gestiona la red Geant) confirmaron que el enlace en cuestión se trataba de un OC-3 no agregado.

Por tanto, la posibilidad de la reordenación por la propia lógica del nodo intermedio parece correcta, ligándola en principio a una mala configuración. Éste no fué el caso: aunque la reordenación se produce en el procesamiento de los paquetes, se comprobó finalmente que la razón era que el modelo de router utilizado en 62.40.103.70 (Juniper M160) planifica los

paquetes repartidos en **cuatro** procesadores (SFMs – Switching and Forwarding Modules) que se dedican a conmutar y encaminar los paquetes entre las diferentes interfaces. En situaciones de congestión como la de nuestro caso de estudio, el retardo introducido por cada procesador es diferente. Nótese que el número de procesadores (4) corresponde con las líneas de tendéncia del retardo de los paquetes.

En este ejemplo hemos podido comprobar la utilidad de la información de la topología de red. Disponer de una información histórica de la misma puede resultar útil más aún cuando puede ser normal querer comprobar información anterior a cambios realizados en la red que no han tenido el rendimiento esperado. Para tener una información topológica suficientemente completa se requiere un buen número de nodos, cuantos más mejor. En este caso concreto, el número de nodos intermedios (routers) promedio entre cualquiera de los extremos de la red es de 15 y el número de nodos desde los que se realizaron pruebas 9.

5 Descubrimiento de topologías y evolución temporal de rutas

El sistema ETOMIC es capaz de ejecutar cualquier software y como tal cualquier aplicación conocida como la herramienta *traceroute*, por ejemplo con el fin de comprobar la topología de la red. En un estudio realizado durante el mes de Marzo de 2005 observando las rutas entre todos los nodos del sistema, las rutas permanecen mayormente estables. Si se considera como observación las ejecuciones periódicas de traceroutes entre un par de nodos durante un día, se han apreciado cambios en las rutas en menos del 7% de las observaciones. La mayoría de estos casos corresponden a cambios en la ruta que han durado unos minutos, para después volver al camino original. Estos cambios transitorios vienen acompañados de perdidas de los paquetes que envía traceroute en un porcentaje apreciable de los casos.

En la Fig. 9 se muestran los RTT mínimos de los saltos desde un nodo situado en Estocolmo (Suecia) hasta otro situado en Jerusalem (Israel). Se ha escogido esta gráfica porque es una de las que tiene más casos de cambios de ruta. En la gráfica se han representado medidas separadas entre si por al menos 24 horas, junto con los cambios que se han detectado. Cuando hay un cambio en una ruta se muestran los saltos nuevos que tiene la ruta. Por ejemplo, se observan cambios momentaneos en la ruta los días 1,3,9 y 10 de Marzo. En la gráfica se ve como aparecen los saltos nuevos de la ruta, representados por un simbolo *'+'* recuadrado y por una *'x'*. El día 12 hubo un cambio de ruta que se mantuvoS hasta el día 23. En ese periodo de tiempo también se observan cambios de ruta de corta duración, especialmente entre los días 12 y 17. Los últimos días del mes, se vuelve a cambiar la ruta por la misma que al principio del mes, que se mantiene estable hasta el final de la medida.

En la Fig. 10 se muestra el grafo de interconexión entre todos los nodos de Evergrow, o dicho de otra forma, todos los enlaces y caminos que se pueden monitorizar en un instante determinado. Esta información de topología del sistema es muy útil para diseñar experimentos por parte de los investigadores. Funcionalidades de este tipo se van a ir incorporando al interfaz web del CMS sucesivamente.

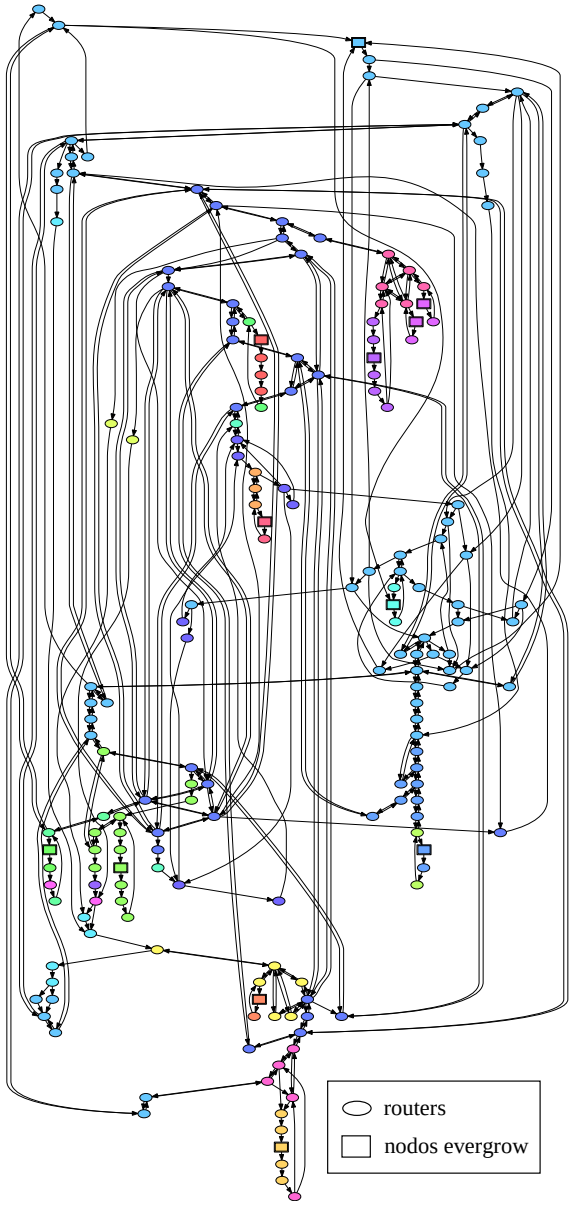


Figura 10: Grafo dirigido que representa los caminos entre todos los nodos de Evergrow. Los diferentes niveles de gris representan sistemas autónomos (AS).

6 Conclusiones

La plataforma ETOMIC provee de una infraestructura de monitorización de altas prestaciones, útil para realizar monitorización activa y pasiva entre nodos distribuidos por toda Europa. Las características de alta precisión del sistema abren la posibilidad

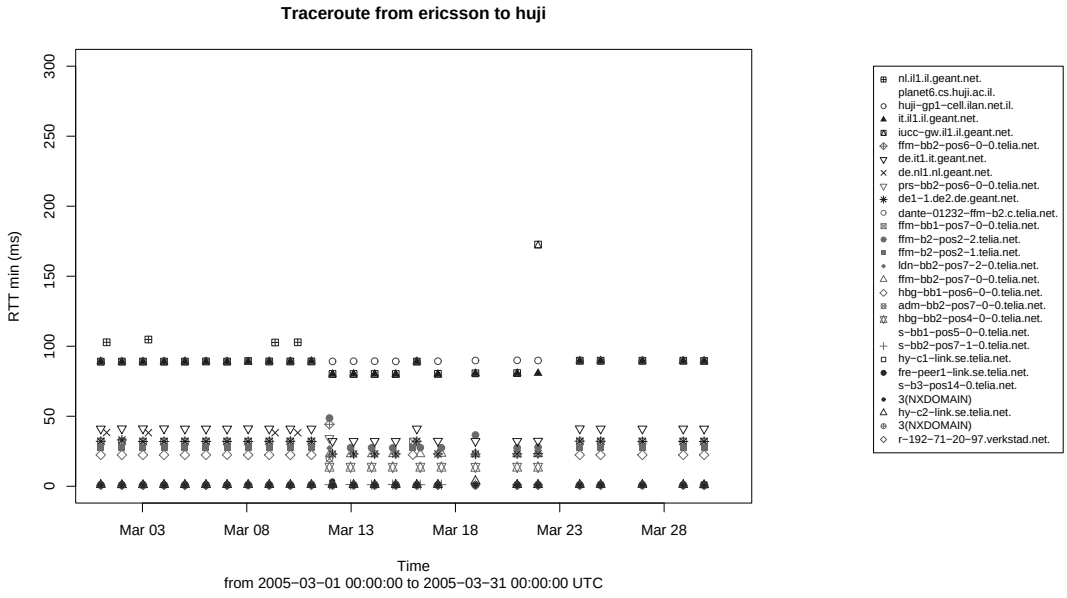


Figura 9: Cambio de rutas a lo largo del mes de Marzo de 2005

de realizar nuevas medidas activas y proponer herramientas que extraigan mayor información de la red. En este trabajo se ha estudiado el estado de enlaces de la Internet europea en base a algunas mediciones activas, observándose fenómenos de interés como la detección de mecanismos de retardo variable introducido por routers debido a su arquitectura de multiconmutación paralela. También se han presentado medidas de retardo en un sólo sentido, y algunas conclusiones que se pueden obtener directamente de ese parámetro como la detección de situaciones de congestión o estimación del ancho de banda del cuello de botella del camino entre dos nodos.

Referencias

[1] A. Pasztor and D. Veitch. On the Scope of End-to-End Probing Methods. *IEEE Communications Letters*, 6(11):509–511, 2002.

[2] E.Magaña, D.Morató, M.Izal, J.Aracil, F.Naranjo, F.Astiz, U.Alonso, and et al. The european traffic observatory measurement infraestructure (ETOMIC). In *IEEE International Workshop on IP Operations & Management (IPOM 2004)*, Beijing, China, October 2004.

[3] D.Morató, E.Magaña, M.Izal, J.Aracil, F.Naranjo, F.Astiz, U.Alonso, and et al. The european traffic observatory measurement infraestructure (ETOMIC): A testbed for universal active and passive measurements. In *Testbeds and Research Infrastructures for the DEvelopment of NeTworks and COMmunities (Tridentcom 2005). Best testbed award*, Trento, Italy, February 2005.

[4] Endace Measurement Systems. http://www.endace.com.

[5] Darryl Veitch, Satish Babu, and Attila Pasztor. Robust synchronization of software clocks across the internet. In *IMC '04: Proceedings of the 4th ACM SIGCOMM conference on Internet measurement*, pages 219–232, New York, NY, USA, 2004. ACM Press.

[6] G. Almes, S. Kalidini, and M. Zekauskas. RFC2679: A one-way delay metric for IPPM, September 1997.

[7] Y. Kitatsuji, M.Tsuru, S. Katsuno, and Y. Oie. Uselfulness of precise time-stamping for exposing network characteristics on high-speed links. volume 5598, pages 163–170, October 2004.

[8] O. Gurewitz and M. Sidi. Estimating one-way delays from cyclic-path delay measurements. In *Proceedings of Infocom, 2002*, 2002.

[9] C. Dovrolis, P. Ramanathan, and D. Moore. Packet dispersion techniques and capacity estimation. *IEEE/ACM Transactions on Networking*, 12(6), 2004.

[10] B. Braden, D. Clark, and et al. RFC2309: Recommendations on queue management and congestion avoidance in the internet, April 1998.

[11] Richard Nelson, Daniel Lawson, and Perry Lorier. Analysis of long duration traces. *SIGCOMM Comput. Commun. Rev.*, 35(1):45–52, 2005.

[12] Geant. Network topology. URL, December 2004. http://www.geant2.net/upload/pdf/GEANT_Topology_12-2004.pdf.

Modelo de análisis y gestión de la calidad de los servicios de telecomunicación: caso de aplicación servicio web (HTTP)

Fidel Liberal, Armando Ferro, Jose Luis Jodra, Eva Ibarrola
Departamento de Electrónica y Telecomunicaciones. UPV/EHU
ESI de Bilbao. C/ Alameda Urquijo S/N.
48013 – Bilbao (Vizcaya)
Teléfono: 946 01 41 29 Fax: 94 601 42 59
E-mail: {jtplimaf, jtpfevaa, jtpjoluj, jtpibara}@bi.ehu.es

Abstract. *This paper presents a comprehensive model intended to analyse quality of service in telecommunications services. Many works have been published in this area, both from a technical viewpoint as well as taking into consideration subjective concerns. However, these works have not resulted in a unique methodology to assess the experienced quality. While most of the studies consider the quality of service from biased perspectives, we try to analyse quality of service as a general gauge of final users' satisfaction. The proposed model allows us to estimate the quality experienced by end users, while offering detailed analysis regarding the different agents involved in the service provision. Once we overview the most significant elements of the model, an in-depth analytical study is detailed. Finally, we illustrate a practical study for HTTP service in order to validate the theoretical model.*

1 Introducción

En los últimos tiempos, al revisar muchos trabajos relacionados con la "calidad de servicio", se observa que comienzan analizando las insuficiencias de las redes actuales, justificando así la necesidad de desarrollar y desplegar arquitecturas y mecanismos de análisis, provisión y gestión de la calidad en las redes de datos. Así, en la mayoría de los casos se apuntaría al carácter *Best Effort* de las redes basadas en TCP/IP, como la causa de todos los males de la QoS, responsable de imposibilitar en la práctica la garantía a gran escala de unos mínimos de calidad. Evidentemente, no sería la única inculpada. Aparecerían otros responsables, como el carácter descentralizado de Internet, la falta de modelos válidos de comercialización de servicios diferenciados entre proveedores, la mala escalabilidad de los mecanismos de provisión de QoS y así un largo etcétera. En base a esas deficiencias se expondría qué aspectos de las mismas se pretendían subsanar y comenzaría así el desarrollo de la arquitectura o mecanismos propuestos. En definitiva, el estudio se centraría únicamente en aspectos técnicos de la prestación de los servicios de telecomunicación y en cómo conseguir una mejora de los parámetros asociados.

El trabajo aquí recogido pretendía partir de un enfoque más generalista que el correspondiente a las siglas QoS. Esto es, no se abordó el problema de la calidad de servicio desde un prisma exclusivamente técnico, sino que se aspiraba a combinar las dos facetas, objetiva y subjetiva, del término calidad en la prestación de servicios de telecomunicación en general.

Partiendo de la doble naturaleza de la calidad parece lógico pensar que un enfoque que se centre exclusivamente en una de las facetas será cuando menos insuficiente, en tanto en cuanto cualquiera de las dos realidades viene determinada o al menos influenciada por la otra. Sin embargo, en general se ha hecho un énfasis mucho mayor en el desarrollo de modelos, arquitecturas, protocolos y herramientas capaces de gestionar los parámetros técnicos de rendimiento de las redes. Las contribuciones orientadas a estudiar la calidad desde el punto de vista de la percepción que el usuario tiene del servicio son mucho menores en número, importancia e inclusión en los estándares. Aunque parezca un contrasentido, en muchos casos se ha obviado el fin, la calidad en su totalidad, y se ha limitado el estudio al análisis y gestión de una lista de parámetros exclusivamente técnicos.

Ante esta situación se ha llevado a cabo un análisis de las diferentes necesidades y carencias detectadas en los actuales estudios y trabajos de QoS. A partir de ese estudio se ha propuesto un modelo general de análisis y gestión de la calidad de servicio, orientado a contemplar las facetas objetivas y subjetivas de forma analítica. En base a ese modelo general se ha propuesto una formulación que guíe la aplicación del modelo y se ha comprobado su utilidad aplicándolo en varios casos de estudio.

El artículo está estructurado de forma similar al estudio realizado: En el apartado 2 se resumen las carencias detectadas en el análisis de iniciativas. En los apartado 3 y 4 se describe brevemente el modelo propuesto y la formulación y notación derivadas del mismo. A continuación se define el caso de aplicación y los resultados del mismo (apartado 5) y se muestran finalmente las conclusiones (apartado 6).

2 Análisis de iniciativas: carencias detectadas

La mayoría de los estudios, iniciativas, normas técnicas y protocolos relativos a la QoS tradicional se centran únicamente en los aspectos objetivos. Esto es, inicialmente, enumeran un conjunto de los parámetros técnicos más comunes, que teóricamente representan la calidad asociada a la transmisión de información. A partir de esa lista cerrada, proponen un nuevo sistema encargado de analizar o gestionar los elementos de red en base a esos parámetros, buscando una optimización de los protocolos o sistemas actuales. Esa optimización se justifica en base a la mejora de dichos parámetros.

Sin embargo, en ningún momento se cuestiona si esa mejora de los parámetros técnicos de operación de los equipos de red redunda en una mejora real para los usuarios, o si se va a traducir en una ventaja competitiva de los proveedores respecto a sus competidores. Es decir, se obvian las características no técnicas de la calidad en sus dos variantes: por un lado la calidad que realmente perciben los usuarios, objetivo último de la prestación de servicios; por otro lado, la calidad de servicio vista como una herramienta más de gestión dentro de un proceso productivo, asociada a los procedimientos de calidad de la empresa proveedora, con limitaciones en cuanto a costes y condicionada por el retorno de la inversión.

2.1 Ausencia de metodología del análisis previo

En algunos casos **[1]**, **[2]** la lista inicial de parámetros sujetos a estudio no se limita a un conjunto de parámetros técnicos sin discusión previa. En lugar de eso, se trata de tener en cuenta cuáles de esos parámetros pueden resultar más relevantes. Así, generalmente mediante encuestas, se consigue una lista de parámetros de partida simplificada y de mayor valor representativo.

Sin embargo, esa fase inicial de evaluación de la importancia de los parámetros técnicos en general no atiende a una metodología concreta. Con cierto criterio, se consulta a los agentes interesados en la cuestión de la calidad (tales como asociaciones de fabricantes, de operadores, de usuarios, organismos de regulación…). Sin embargo, el tratamiento estadístico e incluso el procedimiento de realización de consultas no atiende a una metodología concreta recogida en el propio estudio. Así, por ejemplo, la lista definitiva de parámetros no hace referencia a la posible importancia relativa entre ellas o peso de los diferentes parámetros.

2.2 Estudio subjetivo exclusivamente en la fase inicial

Como se ha comentado, son pocos los estudios que parten de una lista de parámetros que tenga en cuenta la faceta subjetiva de la calidad. Desgraciadamente, la subjetividad en muchos casos se tiene únicamente en cuenta en esa fase inicial. Una vez se confecciona esa lista de parámetros iniciales, el resto del estudio se centra únicamente en proponer mecanismo de mejora de esos parámetros.

De la lista inicial se deriva que existe una relación entre los parámetros técnicos y las sensaciones de calidad de los usuarios. Sin embargo la mayoría de los estudios no ahondan en esa relación. Simplemente, aprovechan la lista inicial de parámetros y no tratan de establecer la relación inversa entre la mejora técnica conseguida y la mejora subjetiva asociada.

2.3 Estudios subjetivos específicos

Los trabajos encaminados a analizar la relación entre las componentes subjetivas y objetivas de la calidad en general están enfocados a factores muy concretos. Esto es, fijando el resto de posibles variables, se centran en valorar cómo afectan las variaciones de ciertos parámetros de red sobre un factor concreto de la percepción de un usuario. Así, en **[3]** se centran en el análisis de la percepción de la calidad de VoIP en función de las pérdidas, el tamaño de los paquetes y la duración de las ráfagas. Lo mismo sucede en **[4]** y **[5]** para la transmisión de video, separando el análisis en función de cómo afectan las pérdidas de la transmisión a cada frame.

Sin embargo, la formulación, los procedimientos de validación y los modelos propuestos están limitados a esos parámetros concretos que afectan a una determinada percepción de un servicio. Por tanto, son difícilmente extrapolables a otras características del propio servicio o no reflejan posibles dependencias adicionales respecto a otros parámetros de red.

2.4 Análisis centrados en un sólo servicio

Existen precedentes de estudios orientadas a establecer las relaciones completas entre la satisfacción de un usuario respecto a un servicio concreto. Así, en **[6]** se realiza un análisis del rendimiento asociado a los protocolos de correo. En otros, como la aplicación del modelo de Kano realizada en **[7]**, se estudia la satisfacción de los usuarios en base a una serie de parámetros subjetivos específicos del servicio. Desgraciadamente, una vez más, los elementos considerados en esos modelos, la formulación de las interacciones y las conclusiones de esos estudios están circunscritas únicamente al ámbito del servicio sujeto de análisis.

2.5 Estudios extremo a extremo, no analíticos

Existen numerosos estudios encaminados a analizar cómo afectan los diferentes tramos de la comunicación a los parámetros de QoS. Esos estudios incluyen los referidos a ingeniería de tráfico y a

modelos y sistemas de simulación de red. En definitiva, las interacciones entre elementos encargados de la provisión de un servicio de telecomunicación son de sobra conocidas, han sido modeladas en los últimos tiempos mediante diversos instrumentos matemáticos y hay disponibles numerosas herramientas de simulación capaces de caracterizarlas fielmente.

Sin embargo, esos estudios relativos al funcionamiento interno de las redes y sus consecuencias en los parámetros finales de las transmisiones no han sido incluidos en los estudios generales de QoS. En lugar de eso, la mayoría de normas y estudios técnicos reducen la visión de la QoS al estudio extremo a extremo. Esto es, analizan el resultado último en términos de calidad en la provisión del servicio, pero no las causas del mismo. Además, no proporcionan mecanismos capaces de identificar los posibles responsables de una degradación del servicio.

3 Descripción del modelo ampliado

El modelo que se describe a continuación trata de dar solución a las carencias recogidas en el apartado anterior. Así, su objetivo fundamental es proporcionar una formulación y una metodología que aborden la problemática de la calidad en todas sus vertientes. La idea fundamental del modelo fue presentada previamente en **[8]** y se ilustra en la Fig. 1. En ese trabajo previo se presentaba una versión preliminar del modelo, en la que no se había desarrollado la parte correspondiente a la formulación matemática ni profundizado en la aplicación práctica.

El funcionamiento y aplicación del modelo se deducen del propio proceso que lleva a la definición de los diferentes elementos del mismo y sus interacciones:

Así, dado que pretender cuantificar la calidad global que obtiene un usuario no parece razonable es preciso referirse de forma separada a cada uno de los servicios finales que se ponen a su disposición. Por tanto, uno de los elementos que debería quedar reflejado en el modelo son los posibles **servicios individuales** que componen el servicio global prestado.

La primera pregunta que va a guiar el desarrollo del modelo es simple, ¿qué hace pensar a un usuario que un determinado servicio es de mayor o menor calidad?. En definitiva, puesto que se trataba de desarrollar un modelo que permita, por ejemplo, verificar que el usuario está obteniendo lo que desea, surge la pregunta, ¿qué es lo que el usuario desea de un servicio?. Cualquier usuario común de Internet puede responder a esa pregunta inmediatamente, y es que, desde su punto de vista, la calidad depende de sus propias percepciones del servicio.

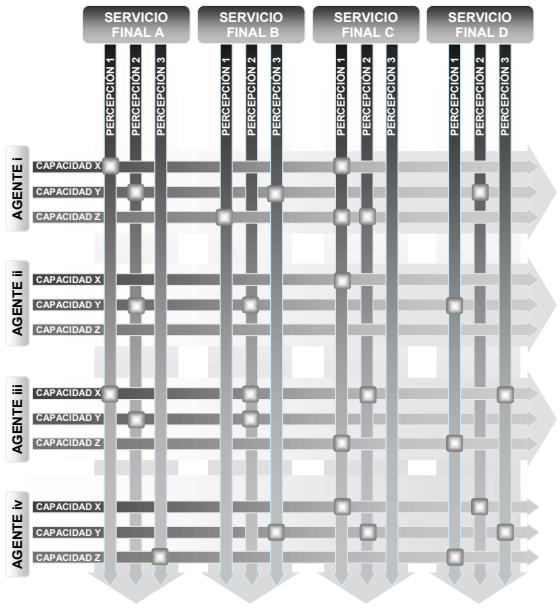


Figura 1. Estructura del modelo general.

Entre las respuestas se podría encontrar: "q*ue los ficheros se bajen rápido*", "*que no se corte la conexión*", "*que las radios on-line no se entrecorten*", etc. Se trata por tanto de valoraciones que para el usuario son más o menos subjetivas. Esas valoraciones se realizan sobre las **percepciones del usuario** y, por tanto, ése va a ser el siguiente elemento del modelo, puesto que son determinantes para que refleje claramente las preferencias de los usuarios.

En este punto ya se han descrito de los elementos verticales del modelo: los usuarios evalúan la calidad global que tiene para ellos el servicio de telecomunicación (por ejemplo el servicio general de "acceso a Internet"), a través de los diferentes servicios finales (DNS, WWW, FTP, mail, video, voz…) por medio de percepciones tales como disponibilidad del servicio requerido, tiempo de respuesta, etc...

En el otro eje tenemos a los **agentes**, que contribuyen a que los usuarios acaben obteniendo el servicio requerido y con la calidad deseada. Evidentemente, cada uno de los agentes que intervienen en la cadena ejerce un determinado rol, que puede afectar o no a ciertos servicios, con un peso concreto, en función de sus características. Cada agente, por tanto, tiene una determinada responsabilidad con respecto a un servicio, en cuanto a que tiene que cumplir su función, además con unas garantías de calidad mínimas. Es decir, el modelo debe permitir identificar de forma sencilla cuáles son las funciones (y por tanto las responsabilidades) de cada uno de los agentes que intervienen en la prestación de los servicios. Para ello, el modelo va a representar cada una de las **funciones genéricas o capacidades** que proporciona un agente.

Se completa por tanto la parte horizontal del modelo, considerando los diferentes agentes o entidades que

intervienen en algún punto del servicio final a usuario con sus respectivas funciones (conectividad, autenticación, prestación de contenidos, etc…).

Queda pendiente una de las cuestiones fundamentales que pretendía resolver el modelo: identificar cómo depende la sensación de calidad que acaba obteniendo el usuario de un determinado servicio de las funcionalidades de que es responsable cada uno de los agentes que intervienen en ese servicio.

El modelo nos proporciona la solución en los **puntos de cruce o "matches"** de los elementos verticales. Cada uno de esos "matches" indica que esa percepción del usuario (el elemento vertical) se ve afectada por la función concreta del agente. Evidentemente es preciso obtener más información que la simple influencia o no. Por ejemplo, se debe categorizar cómo afecta a la percepción X la función Y. Para ello, se define que uno o varios **indicadores** asociados a cada "match" van a determinar de forma cuantitativa cómo se ve afectada la percepción. Con ello se da respuesta a la primera de las preguntas planteadas. En general esos indicadores podrían ser cuantificados por el agente correspondiente mediante medidas.

Sin embargo, existe una cuestión adicional relativa a la posibilidad de establecer la correspondencia entre los parámetros internos de los agentes prestatarios y la percepción de la calidad de los usuarios finales. De esta manera los agentes podrían variar dichos parámetros para mejorar la percepción. Debe quedar claro que, aún cuando los indicadores revelan una dependencia con esa percepción, son en general magnitudes medibles pero **no necesariamente parámetros internos modificables por los agentes**, dado que muchas veces o dependerán de la conjunción de varios de ellos o ni siquiera existirá una correspondencia inmediata. Para ilustrar esa diferencia de matiz, asociado a cada **indicador** se deberá representar en el modelo una serie de **parámetros internos** del agente asociados, debiendo definirse la función que relaciona el indicador con los parámetros.

El modelo propuesto, siguiendo esa estructura matricial, permite relacionar el objetivo final, la sensación de calidad que acaban obteniendo los usuarios a partir de sus percepciones, en función de las causas últimas, los parámetros internos que gestiona cada uno de los agentes que interviene en la prestación de los diversos servicios.

De este modo la aplicación del modelo proporciona una metodología estricta de identificación de los elementos y de las contribuciones de cada agente o grupo de agentes a la satisfacción final del usuario. Eso permite, de forma analítica identificar los responsables de la calidad que perciben los usuarios, como por ejemplo condiciones de cuello de botellas que afectan a la calidad y realizar un despliegue inteligente de recursos.

La utilización del concepto de percepción y la especificación de una metodología y procedimientos de aplicación analíticos suponen además un avance respecto a otros modelos de gestión de calidad matriciales, como el modelo tradicional de gestión de la calidad total que integra QFD **[9]** y las tablas más recientes de indicadores de calidad en servicios de telecomunicación propuestas por Oodan en **[10]** y recogidas en la recomendación G.1000 **[11]**.

4 Desarrollo analítico del modelo

En un modelo general, se considera un número discreto M de servicios a evaluar, proporcionados por un número de agentes N.

Cada agente es responsable de prestar un determinado conjunto de funcionalidades. Para evitar la ambigüedad con el concepto de función que se va a definir más adelante, denominaremos capacidades a esas funciones de cada agente. De este modo, el agente n-ésimo es responsable de prestar un conjunto de capacidades C_n.

Entre esas capacidades, dependiendo del agente encontraríamos la capacidad de visualización de la información, de transmisión, de calidad del contenido, etc…

Por otra parte nos encontramos, que para un determinado servicio "m", un usuario obtiene una mayor o menor satisfacción en base a un conjunto P_m de percepciones. Esas percepciones incluyen la *fiabilidad*, la *velocidad* del servicio, la *calidad* del sonido, etc… Cada percepción tendrá un determinado peso en la sensación de calidad final que tiene el usuario para ese servicio. Esos pesos se calculan aplicando la metodología AHP de forma similar a la recogida en **[12]** y **[13]**. Esa metodología, utilizada para el análisis multicriterio, proporciona mecanismos para ponderar sistemáticamente la importancia relativa de cada percepción. Además, se utilizará análogamente AHP para obtener la satisfacción global del usuario respecto a varios servicios.

La forma en la que está estructurado el modelo permite reutilizar los estudios exclusivamente subjetivos descritos en el apartado 2.3 a la hora de estudiar los factores de los que depende una determinada percepción. Así cada percepción dependerá de los denominados Factores Globales de Valoración (***FGV***), multidimensionales expresados, para la percepción mt-ésima en forma vector como $\vec{G}^{mt}$.

Cada una de las dimensiones de ese vector corresponde a uno de los "puntos de entrada" o ejes

de las gráficas de percepción de calidad que suelen acompañar a los estudios subjetivos.

La valoración final de la percepción se obtendrá a partir de la función de valoración V^{mt}, que será específica de esa percepción y que toma los vectores ***FGV*** como parámetro. Debe quedar claro que los ***FGV*** no son intrínsecamente aspectos relacionados directamente con los parámetros de funcionamiento de la red, sino más bien el resultado de los mismos. Por ejemplo: la velocidad de descarga completa de una página web, el *framerate* efectivo obtenido, características de la codificación de la fuente, etc… Es en esos dos puntos donde se identifica el nexo entre las variantes subjetivas y objetivas de la calidad. Por un lado la función de valoración permite medir la percepción subjetiva a partir de unos determinados factores. Por otro lado esos ***FGV*** dependerán a su vez del resultado del funcionamiento del servicio, expresado en términos objetivos y cuantitativos.

Para expresar esa dependencia se introduce la función de parametrización (***FP***) que es la encargada de obtener el efecto final que obtiene el usuario derivado del funcionamiento y las características objetivas del servicio y agentes. Esa función en algunos casos podrá deducirse de forma analítica, pero en general se podrán utilizar modelos de simulación y/o medidas reales.

Para medir el rendimiento objetivo del servicio, se parte de los parámetros que cada agente n-ésimo es capaz de gestionar de cara a proporcionar la capacidad s-ésima. Esos parámetros incluyen la capacidad de los enlaces, la topología de la red, las características de los equipos de conmutación, etc.

En base a esos parámetros internos (y a través de la denominada función de rendimiento $f_{ns\ j}^{mt}$), el agente proporciona una determinada capacidad con ciertas características de funcionamiento. Esas características de funcionamiento pueden ser de muy distinta naturaleza, en base a lo que en el modelo se denominan categorías y entre ellas podemos encontrar los parámetros tradicionales de QoS como throughput en ambos sentidos, pérdidas, retardos, *jitter*... Esas categorías que reflejan las características de la capacidad que pueden influir en las diferentes percepciones suponen que el concepto de indicador introducido en la descripción general deba ser multidimensional:

$$\vec{I}_{ns}^{mt} = \left(i_{ns\ 1}^{mt}, i_{ns\ 2}^{mt}, \ldots, i_{ns\ j}^{mt}, \ldots, i_{ns\ D^{mt}}^{mt} \right) \quad \textbf{(1)}$$

Ese indicador corresponde a un punto de cruce de la percepción mt para la capacidad ns. El conjunto de indicadores vector correspondientes a todos los cruces entre una percepción mt y un agente viene expresado de forma matricial mediante $\left[\mathbf{I}_{\mathbf{n}}^{\mathbf{mt}}\right]$.

En muchos casos no habrá una dependencia clara y analítica entre los parámetros de los agentes y los indicadores, obteniéndose estos mediante medidas o modelos de simulación.

Esos indicadores se pueden agrupar para identificar las contribuciones de cada agente. Teniendo en cuenta que para una percepción concreta mt los diferentes indicadores se expresan en forma de un vector de contribución de la misma longitud, se va a definir un nuevo concepto denominado Indicador Ponderado de Categoría (***IPC***). Este Indicador Ponderado de Categoría $L_{n\ i}^{mt}$ se define como un indicador real que engloba las contribuciones de un agente n a la categoría i y el vector $\vec{L}_{n}^{mt}$ como aquél cuyas componentes corresponden a los IPCs de un agente a las diversas categorías. Ese procedimiento mediante el cual a partir de la matriz $\left[\mathbf{I}_{\mathbf{n}}^{\mathbf{mt}}\right]$ se obtiene el vector $\vec{L}_{n}^{mt}$ se va a denominar proceso de ponderación local (***PPL***) y va a tener asociada una función h, de ponderación local (***FPL***) y una matriz de pesos $\left[\mathbf{W}_{\mathbf{L}}\right]$ que ilustra por ejemplo el hecho de que no todo los flujos se vean afectados de la misma forma. El proceso se muestra en la Fig. 2.

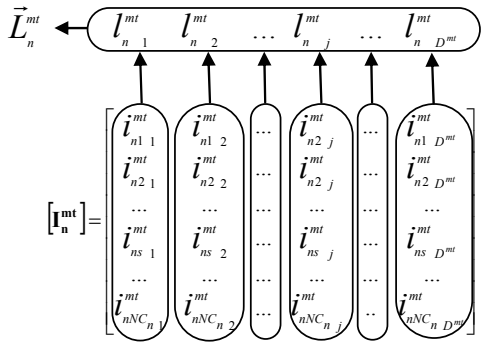


Figura 2. Función de ponderación local.

Del mismo modo que se agrupan las contribuciones de un agente a la misma categoría, se va a realizar un proceso análogo, denominado proceso de ponderación global (***PPG***), para conseguir el vector ponderado general de contribuciones para una determinada percepción "mt". Ese proceso permite obtener el resultado final de las contribuciones de los diversos agentes. Con esta operación se obtendría los indicadores globales que van a permitir obtener los ***FGV***.

Ese proceso de ponderación global, puede ir desde una simple aproximación consistente por ejemplo en considerar el retardo total la suma de los retardos de cada agente y el throughput efectivo el mínimo de los agentes, hasta la utilización de herramientas de simulación o análisis de ingeniería de tráfico.

El proceso completo se describe de forma resumida en la siguiente figura:

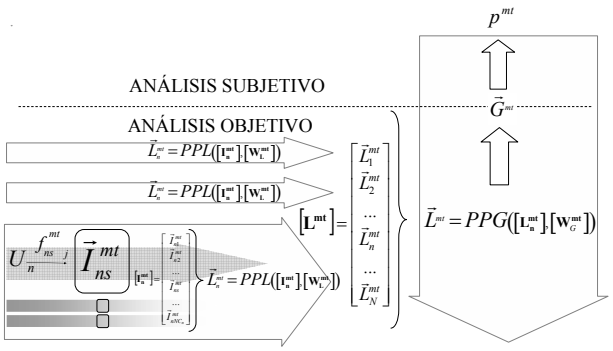


Figura 3. Proceso completo de cálculo de factores en el modelo.

5 Caso de aplicación: servicio web

El escenario de aplicación del modelo consiste en el análisis de la satisfacción de un usuario debida al servicio web. Para ello se propone un modelo de red típico compuesto por los agentes: Plataforma del usuario, red de acceso, ISP, carrier inter-ISPs y proveedor del servicio.

Se pretende analizar la variación de la percepción del usuario para el servicio web bajo unas condiciones técnicas de los agentes concretas y estudiar su variación en función de diferentes velocidades de acceso y su dependencia respecto a los diversos agentes.

Puesto que el servicio a analizar es el web el primer paso consistiría en analizar la percepción subjetiva de calidad que tienen los usuarios. Para ello existen numerosos estudios que cubren parámetros no funcionales (como el descrito en **[14]**), relacionados con el contenido de las páginas web, tales como accesibilidad, seguridad, facilidad de uso, etc... Sin embargo, esos parámetro dependen casi exclusivamente del contenido y no de los parámetros de transmisión, por lo que no serán tenidos en cuenta en este análisis.

La percepción de calidad del servicio web viene expresada en términos comunes como la "velocidad de descarga de páginas" y ha sido analizada en diferentes estudios. Así, en **[15]** se describe la utilización de unos cálculos matemáticos denominados *Fun Factors* como forma de evaluar esa percepción. En **[16]** se analizan las posibles causas de la latencia en el tráfico web. Esa importancia de la "latencia" o tiempo de descarga de la página como factor determinante en la calidad que perciben los usuarios del servicio web viene recogida en **[17]**. Dicho análisis final se completa con análisis más detallados de la percepción de la calidad web en función de la latencia en diversas situaciones. Así, en **[18]** se estudia en detalle cómo varía la tolerancia a la latencia de los usuarios en función del tiempo que ha permanecido éste en el sitio web hasta ese momento.

En este estudio se va a considerar los resultados de **[17]**, según los cuales la percepción de los usuarios acerca de la velocidad de navegación, recogida mediante encuestas en forma de Mean Opinion Score (**MOS**) se aproxima mediante una expresión logarítmica del tiempo de descarga completo de la página (fórmula (**2**)).

Con estas premisas como punto de partida los elementos identificados en el modelo son los siguientes:

- **Percepción:** Velocidad de navegación**.** p^{mt}
- **Valoración:**
  $$MOS = 6 - \log 2(FGV) \quad 1 < MOS < 5 \qquad \textbf{(2)}$$
  con 5 máxima calidad y 1 mala calidad.
- **FGV**: tiempo total de descarga de página**.** $\vec{G}^{mt}$.
- **IPG** $\vec{L}^{mt}$ con categorías: retardo descendente extremo a extremo, retardo ascendente extremo a extremo, velocidad de descarga descendente "efectiva", característica de la página (número de objetos y tamaño de cada uno), retardo DNS total.
- **IPL:** $\vec{L}_n^{mt}$ con categorías retardo descendente de cada agente, retardo ascendente de cada agente, velocidad de descarga descendente "efectiva" en ese agente, característica de la página (número de objetos y tamaño de cada uno).
- **FPG:** como primera aproximación se va a considerar el retardo global como la suma de retardos de cada elemento y el throughput total efectivo (máximo) como el menor de los throughputs de los agentes.

Todos los elementos anteriores son deducibles bien a partir de otros elementos o a partir de medidas en cada uno de los agentes salvo el **FGV**, el tiempo de descarga de página. Para calcular ese factor global de valoración, que va a ser el punto de entrada de la fórmula de la percepción, es preciso llevar a cabo un análisis del funcionamiento del servicio, en este caso del protocolo HTTP y obtener así la función de parametrización.

El funcionamiento más simple de HTTP engloba la opción de persistencia o *keepalive,* solicitando todos los objetos en una misma solicitud (usando *pipelining*) y sin considerar las restricciones impuestas por la ventana TCP (ventana infinita). Como se observa en la Fig. 4, una vez que se descarga la página HTML y se solicitan el resto de objetos HTTP, el servidor envía continuamente los datagramas (sin esperar confirmación) ocupando al completo el ancho de banda disponible. En base a ese análisis tenemos que, despreciando los tiempos de procesamiento y considerando N objetos de tamaño S_i:

- **FP:** $G^{mt} = T_{DNS} + 2 \cdot RTT + T_{MAIN} + \sum_{i=1}^{N} \frac{Si}{BW}$ **(3)**

Si se desea profundizar en el modelado de HTTP, en **[19]** se realiza un estudio detallado del rendimiento de HTTP utilizando diferentes variantes del protocolo

sobre diversos protocolos de transporte. Modelos más complejos de generación de carga y de funcionamiento interno de los clientes pueden consultarse en **[20]**.

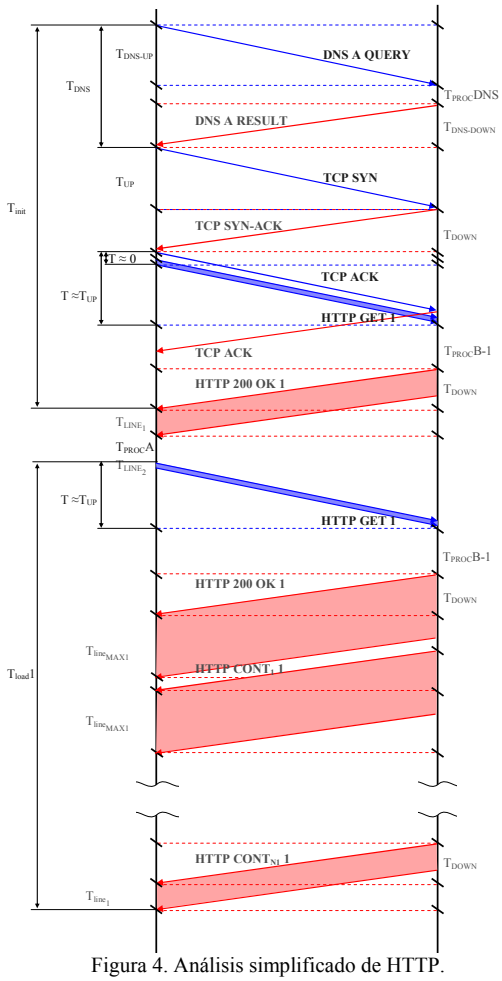


Figura 4. Análisis simplificado de HTTP.

Una vez obtenido de forma analítica la última dependencia ya disponemos de todos los elementos para alimentar el modelo. Para realizar las pruebas se han planteado 5 tipos de página, siguiendo las pautas de **[19]**: small page: 5KB, medium page: 35 KB, large page 100 KB, small cluster 1 pág. Aprox. 6KB y 2 objetos 2,5 KB, large cluster 1 pág. 100 KB y 10 objetos de 25 KB.

En base a esos datos, con unas características de los agentes intermedios dadas se ha procedido a la simulación. Para ello se ha desarrollado un modelo en MATLAB ©, con las simplificaciones teóricas y un modelo "real" en OPNET© usando HTTP 1.1 con *pipelining*. Se ha realizado una serie de simulaciones para diferentes velocidades de la conexión ADSL del usuario obteniendo los resultados de las Figs. 5 y 6. El set A corrobora que la aproximación teórica realizada es aceptable ya que se asemeja al resultado obtenido en la simulación con OPNET©.

En el set B de pruebas realizado en MATLAB© se han modificado las características del resto de elementos. En la Fig. 6 se observa cómo a partir de un determinado throughput en la red de acceso la sensación de calidad de los usuarios no varía.

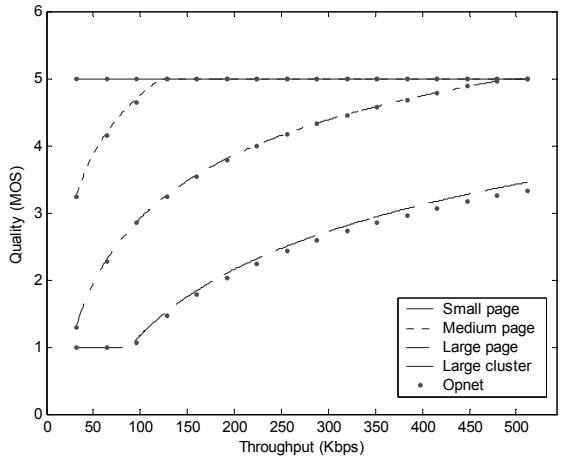


Figura 5. Set A resultados MATLAB© y OPNET©.

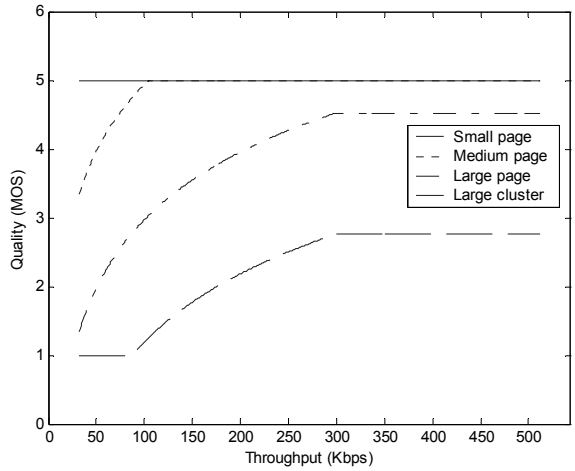


Figura 6. Set B resultados MATLAB©.

Este efecto evidencia que existe un cuello de botella en algún agente que hace inútil el aumento de throughput a partir de ese punto.

Para detectar el responsable de ese efecto se calcula el coeficiente de correlación (la correlación cruzada "normalizada") de la percepción con cada uno de los parámetros de throughput de los agentes obteniéndose, tal como se ve en la Fig. 7, cómo es el enlace del ISP con el carrier el responsable de ese cuello de botella y del estancamiento de la percepción de calidad de los usuarios.

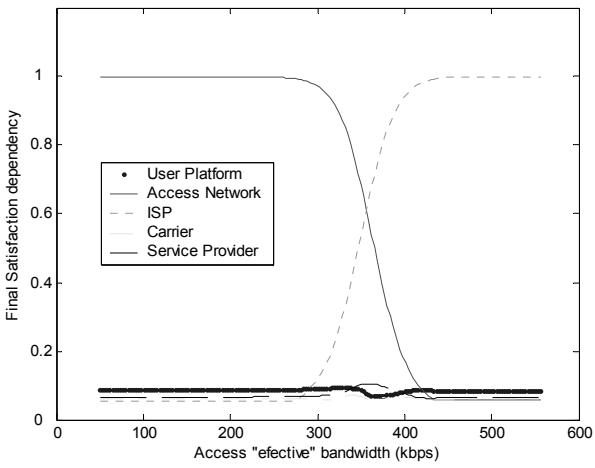


Figura 7. Identificación de cuellos de botella.

6 Conclusiones

El modelo propuesto pretende definir los elementos y la metodología para llevar a cabo un análisis cuantitativo de la calidad que perciben los usuarios de los servicios de telecomunicación. Para ello se ha llevado a cabo una separación de las contribuciones por cada funcionalidad que ofrece cada uno de los agentes que interviene en la prestación del servicio. Esa separación permite reutilizar otros trabajos, medidas de campo o realizar simulaciones para obtener los datos parciales que alimenten el modelo.

Una vez definida la formulación analítica del modelo se ha llevado a cabo un ejercicio de aplicación sobre un caso sencillo, el análisis de la calidad del servicio web en base a la velocidad de navegación, y se ha observado como las aproximaciones teóricas del protocolo se asemejan a los resultados de simulación y cómo el modelo resulta útil de cara a predecir comportamiento y detectar desviaciones como cuellos de botella.

Actualmente se están realizando esfuerzos para comprobar la validez del modelo en entornos más complejos, con diversos servicios y percepciones.

Referencias

[1] Bannock_Consulting, *Quality of Service Parameters for Internet Service Provision*. 2000, European Commision, DG Information Society.

[2] ETSI, *User Group; Quality of Telecom Services; Part 1: Methodology for identification of parameters relevant to the Users*, E.U. Group., Editor. 2002.

[3] E. C. Adrian y Z. Yali, *A simulation-based methodology and tool for automating the modeling and analysis of voice-over-IP perceptual quality.* Performance Evaluation, 2003. **54**(2): p. 147.

[4] J. Klaue, B. Rathke, y A. Wolisz, *EvalVid - A framework for video transmission and quality evaluation.* COMPUTER PERFORMANCE EVALUATION: MODELLING TECHNIQUES AND TOOLS, 2003. **2794**: p. 255-272.

[5] O. Verscheure, P. Frossard, y M. Hamdi, *User-oriented QoS analysis in MPEG-2 video delivery.* REAL-TIME IMAGING, 1999. **5**(5): p. 305-314.

[6] J. Charzinski. *Observations in E-Mail Performance*. en *ITC Specialist Seminar*. 2002. Würzburg, Germany.

[7] P. Zhang y G. von Dran, *User Expectations and Rankings of Quality Factors in Different Web Site Domains.* International Journal of Electronic Commerce, 2001. **6**(2): p. 9.

[8] A. Ferro, E. Ibarrola, y F. Liberal. *Modelo basado en la percepción de los usuarios para la gestión de la calidad de servicio en redes de datos*. en *Jitel 2003*. 2003. Gran Canaria.

[9] Y. Akao, *Quality Function Deployment: Integrating Customer Requirements into Product Design*, ed. P.P. Inc. 1990. 387.

[10] A. Oodan, et al., *Telecommunications Quality of Service Management: from legacy to emerging services*. IEE Telecommunications Series 48, ed. IEE. 2003: IEE.

[11] G.1000[ITU-T], *G.1000 Recommendation*. 2001.

[12] L. D. Bodin, L. A. Gordon, y M. P. Loeb, *Evaluating Information Security Investments Using the Analytic Hierarchy Process.* Communications of the ACM, 2005. **48**(2): p. 79-83.

[13] Ghinea y Magoulas, *Quality Of Service For Perceptual Considerations: An Integrated Perspective.* IEEE International Conference on Multimedia and Expo, 2001: p. 752-755.

[14] S. Barnes y R. Vidgen. *WebQual: An Exploration of Web-site Quality*. en *Proceedings of the Eighth European Conference on Information Systems*. 2000. Viena.

[15] J. Charzinski. *Measured HTTP Performance and Fun Factors*. en *ITC 2001*. 2001. Salvador, BA, Brasil.

[16] A. Habib y M. Abrams. *Analysis of Sources of Latency in Downloading Web Pages*. en *WebNet 2000*. 2000.

[17] R. D. Van der Mei, *"Performance Analysis of Communication Networks" Course.* 2004, Faculty of Science: Vrije Universiteit.

[18] N. Bhatti, A. Bouch, y A. Kuchinsky, *Integrating user-perceived quality into Web server design*, en *Proceedings of the 9th international World Wide Web conference on Computer networks: the international journal of computer and telecommunications netowrking*. 2000, North-Holland Publishing Co. p. 1--16.

[19] O. Riva, J. Saarto, y M. Kojo, *Performance Analysis on HTTP Traffic and Traffic Mixtures with Competing TCP and UDP Flows*. 2004, University of Helsinki - Department of Computer Science.

[20] T. Irnich, *Measuring and modelling WWW traffic characteristics in access networks*, en *Chair of Communication Networks*. 2000, Aachen University of Technology.

Técnicas de agrupamiento vectorial y detección geométrica de anomalías en red

Jesús Díaz-Verdejo, Juan M. Estévez-Tapiador, Pedro García-Teodoro
Departamento de Teoría de la Señal, Telemática y Comunicaciones. Universidad de Granada
ETSI de Ing. Informática. C/ Daniel Saucedo Aranda S/N
18071 - Granada
E-mail: jedv@ugr.es

Abstract *This paper presents some issues concerning N3 -a geometrical based Intrusion Detection System (IDS)- related with its computational performance. Despite its good behavior as an IDS, the scoring of an observed traffic instance by N3 requires its comparison with a set of instances representing the normality model obtained during a training stage. As the size of the training set increases, so do the detection capabilities of N3. Nevertheless, there is a counterpart: a significant increase in computational effort. Our goal is to reduce the size of the normality model and, therefore, the computational requirements of N3, without degrading its detection capabilities. For this purpose, two clustering techniques are proposed and evaluated. The first one is inspired by the well-known k-means algorithm, as k-means is not directly applicable to tackle this problem. The second one is an ad hoc technique developed for this case. Both algorithms allow us to achieve the proposed goal.*

1. Introducción

La gran expansión de Internet, junto con el continuo incremento en el número y gravedad de los incidentes relacionados con la seguridad de la red, hacen de la seguridad en Internet un campo de investigación y desarrollo de enorme transcendencia. Aunque los primeros desarrollos datan de hace varias décadas [1], algunos de los problemas más relevantes continúan pendientes de solución. En este contexto, una de las aproximaciones más estudiadas se basa en la detección de actividades de naturaleza intrusiva, es decir, que violen las políticas de seguridad establecidas en un sistema. Surgen así los denominados Sistemas de Detección de Intrusiones (IDS) [2], cuya finalidad es detectar y alertar sobre cualquier tipo de incidente relacionado con la seguridad. Habitualmente, los IDS se clasifican en dos categorías [3]: los basados en firmas y los basados en anomalías. Los primeros se caracterizan por disponer de una base de datos (*firmas*) con información relativa a los ataques o comportamientos intrusivos conocidos. La segunda categoría, los IDS basados en anomalías, pretenden detectar comportamientos anómalos en el sistema, a los que clasifica como *sospechosos* de ser causados por ataques. La práctica totalidad de los IDS disponibles en la actualidad son del tipo basado en firmas, encontrándose los basados en anomalías en un estado altamente inmaduro [4]. Sin embargo, la investigación en este tipo de IDS es relevante debido a la potencial capacidad de los mismos para detectar ataques no observados previamente y, por tanto, no contenidos en la base de datos de *firmas*. Algunas de las tendencias más recientes se basan en el modelado de protocolos o el análisis de cargas útiles destinadas a servicios específicos [5], [6].

El presente trabajo se centra en un IDS basado en anomalías, desarrollado por los autores, denominado de *Vecino normal más próximo*, N3 (de *Nearest Normal Neighbour*) [7] [8] [9]. Este sistema se basa en la comparación de las secuencias contenidas en las instancias de tráfico procedentes de la red con un conjunto de ellas previamente obtenido durante el entrenamiento del sistema, denominado *modelo de normalidad*. El sistema presenta unos rendimientos altamente satisfactorios. Sin embargo, el algoritmo propuesto presenta una complejidad computacional que depende, directamente, del número de instancias que constituyen el modelo de normalidad. Por este motivo, para la aplicación práctica del sistema, resulta de interés la reducción del número de elementos en el modelo de normalidad sin que ello implique una degradación del rendimiento del sistema.

La aplicación de técnicas de agrupamiento estándar no resulta evidente en el caso que nos ocupa. En el presente trabajo se desarrolla y analiza una modificación del algoritmo de agrupamiento *k-medias*, junto con un método ad hoc para la reducción del número de elementos del modelo de normalidad.

De esta forma, el resto del artículo se estructura como se describe a continuación. En primer lugar, en el Apartado 2 se presenta el sistema N3 para la detección de anomalías, haciéndose énfasis en las necesidades computacionales de dicho sistema. El Apartado 3 describe el escenario experimental utilizado, tanto en lo que se refiere al

entrenamiento como a la evaluación del sistema. En el Apartado 4 se plantean con mayor detalle los problemas asociados al modelo de normalidad y la forma de evaluación considerada, lo que nos llevará a proponer dos algoritmos de agrupamiento específicos, que serán descritos y analizados a continuación. Finalmente, se presentarán las conclusiones y las líneas de trabajo futuras.

2. El IDS de vecino normal más cercano (N3)

El detector de vecino normal más cercano, N3 (de *Nearest Normal Neighbour*) [7] se fundamenta en el modelado del tráfico de red con la finalidad de obtener un modelo del comportamiento normal del mismo. Así, cada observación o evento monitorizado en la red es contrastado y evaluado con el modelo de normalidad disponible (Fig. 1), categorizándose como normal o anómalo en función de una *medida de normalidad* establecida al efecto.

La naturaleza y funcionamiento final del IDS será, obviamente, función de los eventos monitorizados y modelados, junto con la medida o criterio de normalidad utilizado. El objeto de modelado en el sistema N3 son las cargas útiles de las instancias recibidas asociadas a un protocolo particular. Es decir, N3 es un IDS basado en anomalías de protocolos.

El funcionamiento del detector se detalla a continuación. Las cargas útiles extraídas a partir de las unidades de datos del protocolo (H en la Fig. 1) son analizadas y parametrizadas (bloque P). La obtención del modelo (M) se realiza durante una fase de *entrenamiento* del sistema, en la que se utilizan instancias lícitas destinadas al servidor considerado y correspondientes al protocolo a modelar. Estas instancias son procesadas, de acuerdo al procedimiento de extracción y caracterización del modelo (E). Tras esta fase, se dispondrá de un modelo de normalidad del protocolo, $\mathcal{N}_{norm}^{prot}$, compuesto por las cargas útiles observadas. En el modo de detección, las cargas útiles son parametrizadas (módulo P) y clasificadas por el detector (D) a partir del modelo.

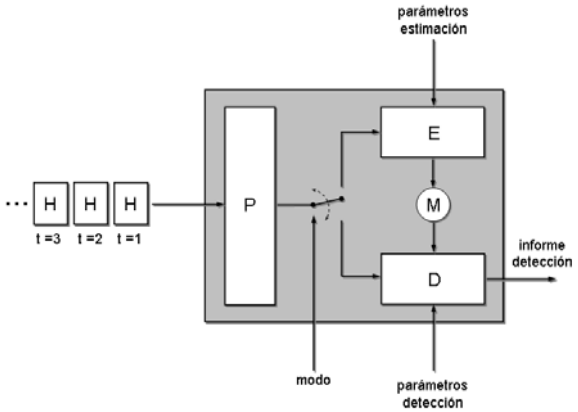


Figura 1: Esquema conceptual de un IDS basado en anomalías.

La detección se realiza a partir de un *índice de anormalidad* de la carga útil, p, $A_s(p)$, obtenida de la comparación de la carga útil a analizar con las existentes en el modelo de normalidad. A este fin, se emplea una técnica de análisis de secuencias cortas, que será descrita más adelante, que establece una medida de distancia, D, entre dos cargas útiles cualesquiera. Por tanto, el índice de anormalidad de la carga útil p se obtiene como la mínima distancia entre dicha carga útil y un elemento del modelo de normalidad:

$$A_s(p) = \min_{\forall q \in \mathcal{N}_{norm}^{prot}} \{D(p,q)\} \tag{1}$$

Una vez obtenido el índice de anormalidad, la clasificación de una carga útil como normal o anómala se puede realizar sin más que establecer un umbral, θ, tal que todas las cargas útiles que superen dicho umbral serán clasificadas como anómalas. Matemáticamente,

$$Clase(p) = \begin{cases} \text{Normal} & si \quad A_s(p) \leq \theta \\ \text{Anómalo} & si \quad A_s(p) > \theta \end{cases} \tag{2}$$

El valor concreto de θ debe ser ajustado experimentalmente mediante algún procedimiento al efecto. Sin embargo, para nuestros fines, dado que lo que se pretende es mostrar la validez de las técnicas propuestas, bastará con evidenciar que existen valores de θ para los que el comportamiento del IDS es el adecuado.

En este trabajo se estudia la aplicación del sistema a instancias de tráfico de los protocolos HTTP y DNS. Estos protocolos transportan elementos con una sintaxis y semántica bien establecidos, definidos en los correspondientes RFC. En concreto, las cargas útiles de HTTP responden a lo establecido en los RFC 2068 y 2396, mientras que las de DNS se definen en el RFC 1034. En ambos casos, dichas cargas útiles corresponden a cadenas de caracteres organizados en secciones y campos, cada uno con un significado y finalidad claros.

2.1. Análisis de secuencias cortas

Como se ha indicado con anterioridad, el detector debe obtener un índice de anormalidad a partir de la evaluación de una serie de distancias entre los elementos del modelo de normalidad y el elemento a categorizar. A este fin es necesario establecer un espacio vectorial asociado a las cargas útiles, de tal forma que se puedan evaluar distancias entre ellas. Para ello se recurre a las técnicas de análisis de secuencias cortas [10], que permiten establecer una representación vectorial de las cadenas de caracteres contenidas en una carga útil y, por ende, de la carga útil en su totalidad. A continuación describiremos brevemente los fundamentos de la técnica empleada. Una descripción más detallada puede encontrarse en [7].

Consideremos un conjunto finito, A, denominado *alfabeto*, de cardinal n, compuesto por elementos denominados *símbolos*. A partir de este conjunto, y fijada una longitud k, es posible definir un *vocabulario* compuesto por todas las posibles concatenaciones de k símbolos. Dado que existen n símbolos, existirán n^k elementos en el vocabulario. Por otra parte, los elementos de este conjunto puede ser ordenados de acuerdo a un criterio arbitrario que permite definir una transformación que representa una secuencia de símbolos del alfabeto de longitud arbitraria, S, mediante un vector de características, $\vec{M}_k(S)$, de acuerdo al siguiente procedimiento: *la componente i-ésima del vector corresponderá al número de veces que la secuencia i-ésima del vocabulario aparece en la secuencia considerada, S.*

Sin pérdida de generalidad, en adelante supondremos que el alfabeto considerado es el código ASCII y que las secuencias a representar corresponden a las cargas útiles de las PDU asociadas al protocolo a modelar.

La transformación propuesta presenta varias características relevantes desde el punto de vista de su aplicación. En primer lugar, dado un vector de características, no existe necesariamente una única secuencia origen que lo genere. Es más, es posible que no exista ninguna secuencia origen. En segundo lugar, se pueden definir las operaciones habituales sobre los vectores de características, generando un espacio vectorial. En particular, se puede definir el producto escalar y, consecuentemente, la norma de un vector, de la forma habitual. Así, dados dos vectores x_1 y x_2, se puede definir una distancia euclídea

$$\begin{aligned} d(x_1, x_2) &= \|x_1 - x_2\| \\ &= \left(\langle x_1, x_1\rangle - 2\langle x_1, x_2\rangle + \langle x_2, x_2\rangle\right)^{\frac{1}{2}} \end{aligned} \tag{3}$$

El producto escalar se puede interpretar como el número de subcadenas comunes que aparecen tanto en x_1 como en x_2. De esta forma, proporciona una indicación de la similitud entre ambos vectores, aunque adolece de referencias temporales en el sentido de que no se considera la ubicación relativa de las subcadenas en cada cadena.

Por tanto, dadas dos cargas útiles, p_1, p_2, es posible definir una distancia, D, entre ellas a partir de la distancia entre su vectores de características,

$$D(p_1, p_2) = d(\vec{M}_k(p_1), \vec{M}_k(p_2)) \tag{4}$$

La distancia entre dos cargas útiles así definida puede ser utilizada por el detector para la clasificación. Sin embargo, es necesario tener en cuenta dos aspectos importantes de cara a su utilización: primero, las distancias se evalúan entre los vectores de características, no entre las cargas útiles directamente; y segundo, la dimensionalidad de los vectores de características es excesiva, incluso para valores reducidos de k. A modo de ejemplo, si se considera el código ASCII como alfabeto y $k = 5$, el vector de características tiene $256^5 = 1099511627776$ dimensiones. Evidentemente, la gestión de estos vectores es computacionalmente muy costosa, incluso inabordable en algunos casos. Por otra parte, la interpretación del producto escalar descrita anteriormente permite la aplicación de técnicas basadas en programación dinámica para su evaluación directa a partir de las cargas útiles [10], lo que permite abordar computacionalmente el problema. Una descripción detallada del algoritmo y sus fundamentos excede los objetivos de este artículo. En cualquier caso, a partir de la Ec. (3), es evidente que puede evaluarse la distancia entre dos cargas útiles a partir del cálculo de 3 productos escalares de acuerdo al algoritmo de programación dinámica. Sin embargo, es importante reseñar que, de esta forma, no se dispondrá en ningún momento de forma explícita de la representación vectorial de las cargas útiles.

2.2. Complejidad computacional

El algoritmo utilizado para la evaluación de las distancias permite obtener el número de subsecuencias de longitud k compartidas por dos secuencias de entrada en $O(n^2)$ operaciones, siendo n la longitud de las secuencias de entrada. Por otra parte, la detección se realiza a partir del índice de anormalidad, de acuerdo a la Ec. (1), lo que requiere la evaluación de las distancias desde cualquier secuencia en el modelo de normalidad a la secuencia a clasificar, así como la comparación de los valores obtenidos a fin de encontrar el mínimo. En otras palabras, se realiza una búsqueda del vecino más cercano en el modelo. Si el modelo se compone de N secuencias, se requerirá evaluar N distancias y realizar $N - 1$ comparaciones. Dado que la complejidad de la evaluación de las distancias es claramente superior a la de las comparaciones, podemos, en primera aproximación, despreciar estas últimas.

Así, el tiempo requerido para determinar si una carga útil dada es anómala o no puede aproximarse mediante la siguiente expresión:

$$t_{detec} \approx C \cdot \tau(n^2) \cdot N \tag{5}$$

donde C es un factor constante que incorpora otras operaciones adicionales requeridas y los detalles dependientes de implementación.

3. Escenario de evaluación

La evaluación del detector propuesto requiere el establecimiento de un escenario de evaluación. Sin embargo, uno de los mayores problemas existentes para el desarrollo de la tecnología IDS es la falta de escenarios y técnicas comunes para su evaluación, complicando enormemente la comparación de los diferentes sistemas existentes [11]. Uno de los que se encuentra disponible para la comunidad investigadora es el Programa de Evaluación de IDS de DARPA [12], que proporciona

Tabla 1: Escenario experimental utilizado.

Conjuntos de datos utilizados				
Nombre	Prot.	Tipo	Descripción	N. elem.
$\mathcal{N}_{norm}^{HTTP}$	HTTP	Normal	Peticiones HTTP para entrenamiento	3262
$\mathcal{N}_{eval}^{HTTP}$	HTTP	Normal	Peticiones HTTP para evaluación	1397
$\mathcal{A}^{HTTP}$	HTTP	Ataques	Variantes de 86 ataques HTTP destinados a *Hume* y *Marx*	119
$\mathcal{N}_{norm}^{DNS}$	DNS	Normal	Peticiones DNS para entrenamiento	46849
$\mathcal{N}_{eval}^{DNS}$	DNS	Normal	Peticiones DNS para evaluación	19934
$\mathcal{A}^{DNS}$	DNS	Ataques	Ataques DNS destinados al servidor DNS	6

trazas de tráfico, convenientemente etiquetadas, obtenido durante varias semanas en una red con varios cientos de ordenadores y conexión a Internet. Aunque el entorno presenta algunas limitaciones (véase [13]), es el más utilizado en la literatura, por lo que será el empleado en este trabajo.

La base de datos de tráfico debe ser particionada en dos bloques: un conjunto de entrenamiento, libre de ataques, y un conjunto de evaluación que, a su vez, debe estar compuesto por tráfico libre de ataques y tráfico correspondiente a ataques.

Los datos de tráfico normal han sido extraídos de la base de datos de Evaluación de IDS DARPA '99. En particular, se han considerado los paquetes correspondientes a las solicitudes de los clientes de los protocolos HTTP y DNS de las semanas 1 y 3, que se encuentran libres de ataques. Los datos extraídos han sido filtrados a fin de eliminar los duplicados existentes, carentes de utilidad para nuestro trabajo. A fin de disponer de datos para el entrenamiento y la evaluación del sistema, se han categorizado estos datos, de forma aleatoria, en dos conjuntos. Asi, el 70 % de los datos se usarán para entrenamiento y el 30 % restante para la evaluación.

Por otra parte, se han generado ataques contra estos servicios, simulando un entorno de red idéntico al anterior. Para ello se ha recopilado la información sobre ataques a los servicios HTTP y DNS descrita en la base de datos arachNIDS [14] y se han implementado programas que los generan.

Finalmente, el escenario experimental consta, para cada protocolo, de 3 conjuntos cuyas características se resumen en la Tabla 1.

4. Reducción de los modelos

La aproximación descrita presenta una alta tasa de detección manteniendo una tasa de falsas alarmas realmente baja [8], lo que posibilitaría su implantación en un entorno real. Sin embargo, presenta una importante limitación en cuanto a su rendimiento computacional. De acuerdo a la Ec. (5), los dos factores determinantes de la complejidad del algoritmo son el cálculo de la distancia y el tamaño del modelo, N.

En el escenario considerado, el tamaño del modelo es de 3262 cargas útiles en el caso de HTTP y de 46849 en el caso de DNS. Por otra parte, el tiempo necesario para realizar el cálculo de las distancias ha sido evaluado durante los experimentos realizados con un procesador Pentium 4 a 2.4 GHz y 1 GB de RAM. Este tiempo presenta un valor medio de 0,00483 ms para HTTP. A partir del valor medio es posible obtener una aproximación al número de peticiones de servicio por segundo que podría procesar el sistema. En el caso de HTTP, el tiempo medio necesario para clasificar una petición es $t_{detec} \approx 0,00483 * 3262 = 15{,}76ms$. Por tanto, se podrán procesar alrededor de 63 peticiones HTTP por segundo.

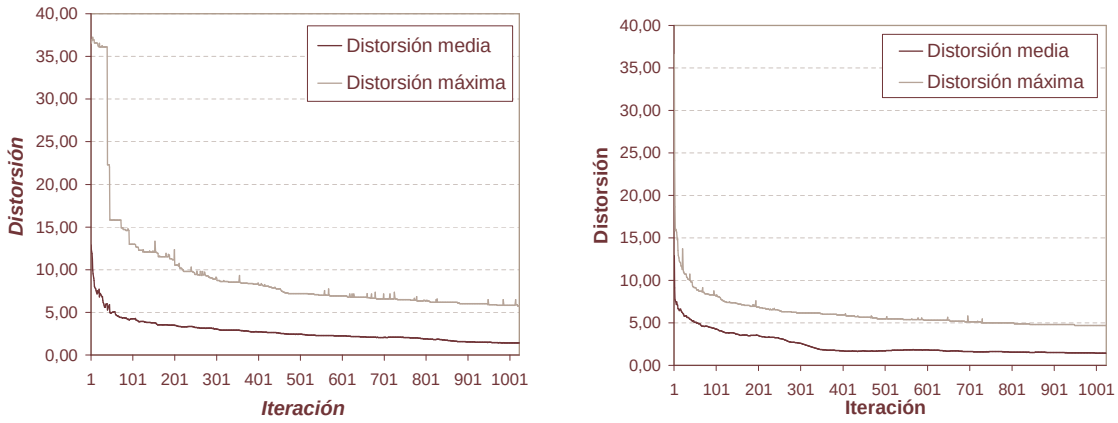
En el caso de DNS, el número de peticiones que se pueden procesar se reduce a alrededor de 5. Obviamente, estos valores resultan inadecuados para una aplicación práctica del sistema, si bien hemos de reseñar que la mayoría de los sistemas de detección de anomalías propuestos hasta la fecha presentan una complejidad computacional claramente superior a ésta.

En cualquier caso, resulta de interés la mejora del rendimiento del sistema en términos computacionales. A este fin se podría reducir la complejidad de la evaluación de la distancia o el tamaño del modelo. El presente trabajo se centra en la segunda solución. Evidentemente, el incremento en el rendimiento computacional será proporcional a la reducción en el tamaño del modelo. Sin embargo, debe realizarse sin degradar el rendimiento del sistema en términos de capacidad de detección.

Por otra parte, la simple inspección de las cargas útiles muestra que, generalmente, existen subcadenas compartidas entre muchas de ellas. Esta afirmación es corroborada por los resultados experimentales obtenidos [8], que avalan la hipótesis de que el tráfico normal se condensa en regiones concretas del espacio de características. Sería posible, por tanto, obtener una representación reducida de estas zonas mediante un número inferior de elementos o representantes. Para ello se pueden utilizar algoritmos y técnicas de agrupamiento.

4.1. Aplicabilidad de los algoritmos de agrupamiento

En una primera aproximación cabría aplicar algoritmos de agrupamiento bien conocidos como podría ser el algoritmo *k-medias* [15]. Sin embargo, algunos aspectos reseñados con anterioridad difi-



a) b)

Figura 2: Evolución de las distorsiones media y máxima de los algoritmos de agrupamiento basados en *k-medias* aplicados a $\mathcal{N}_{norm}^{HTTP}$. a) Algoritmo *k-medias adaptado*. b) Algoritmo $N1$.

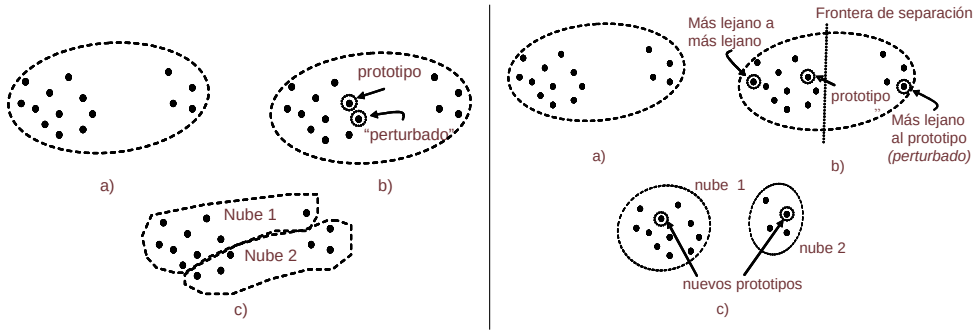


Figura 3: Agrupamiento mediante algoritmos tipo *k-medias*. A la izda., agrupamiento subóptimo cuando se usa como prototipo el vector real más próximo al centro: a) Nube original, b) Representante y su versión distorsionada, c) Regiones resultantes. A la derecha, algoritmo N1: a) Nube original, b) Puntos seleccionados como prototipo y nuevos prototipos preliminares, c) Regiones y prototipos resultantes.

cultan la aplicación directa de estos algoritmos. En primer lugar, no se dispone de las representaciones explícitas de los vectores, lo que invalida cualquier procedimiento que trabaje directamente sobre los mismos. Adicionalmente, incluso en el hipotético caso de disponer de dichos vectores, la aplicación de algoritmos como *k-medias* sería inviable, ya que no podemos invertir la transformación propuesta. Por lo tanto, una vez obtenido un vector *promedio* (*centroide* en la terminología del algoritmo *k-medias*), no sería posible obtener una única secuencia a partir de la que se obtendría el citado vector para su posterior inclusión en el modelo. Idéntica aseveración es aplicable a la versión perturbada del centroide, necesaria para la iteración del algoritmo. Por otra parte, la naturaleza de la transformación puede originar que el vector promedio no corresponda a ninguna secuencia origen, lo que agrava el problema anterior.

Para evitar los problemas antes reseñados se puede utilizar una modificación que evite obtener vectores como promedios o a partir de la perturbación de otros datos. Así, se considera como prototipo (*pseudo-centroide*) de una nube aquella secuencia en el modelo cuya distancia a los restantes miembros sea mínima. La versión perturbada, necesaria para el proceso de división de la nube, será el elemento más cercano al prototipo. De esta forma se garantiza la existencia de elementos en el modelo directamente identificables con los vectores utilizados en el algoritmo. Sin embargo, la aplicación de este procedimiento proporciona unos resultados insatisfactorios al no seleccionar los prototipos adecuadamente. Esto se evidencia tanto en los resultados de detección, no mostrados, como en la evolución de la distorsión promedio y máxima, mostrada en la Fig. 2.a. Como se puede observar, la convergencia no es monótona, como sería de desear, y la distorsión máxima presenta fluctuaciones evidentes. Este último efecto implica que, en algunos casos, la división de una nube no disminuye la distancia media entre sus elementos sino que la aumenta (véase la Fig. 3).

4.2. Algoritmo de agrupamiento N1

Los problemas detectados en la aplicación del algoritmo *k-medias* descrita en el apartado anterior provienen de dos fuentes principales: primero, la división de las nubes se realiza de forma subóptima al considerar puntos reales en lugar de promediados y, segundo, pueden existir puntos que estén más próximos a los representantes de otra nube que al de la suya propia. Para paliar estos problemas, se propone la introducción de dos mo-

Tabla 2: Pseudo-código del algoritmo N1

Inicialización: Se considera una única nube, C_1 a la que pertenecen todos los vectores (N_p) en N_{norm}^{prot}.
Iteración: Repetir hasta que la distorsión media, D_m supere un umbral, θ.
Establecer un conjunto vacío de prototipos $P = \emptyset$
Para cada nube C_i en el espacio N_{norm}^{prot}
Obtener p_{C_i}, la carga útil para la que la distancia media al resto de elementos en la nube es mínima, que será el pseudo-centroide.
Añadir p_{C_i} al conjunto de prototipos, $P \leftarrow P \cup \{p_{C_i}\}$.
Obtener p_{nC_i}, el vector de la nube más distante al pseudo-centroide: $p_{nC_i} = argmax\{d(p_{C_i}, q)\}\forall q \in C_i$. La distancia máxima es el *radio* de la nube.
Seleccionar el cluster C_r con el radio máximo.
Obtener p_{fC_r}, el vector de dicha nube más distante a p_{nC_r}: $p_{fC_r} = argmax\{d(p_{nC_r}, q)\}\forall q \in C_r$
Dividir la nube en dos reasignando todos sus puntos a p_{nC_r} o a p_{fC_r} en función de un criterio de mínima distancia.
Obtener los nuevos pseudo-centroides, p_{1C_r} y p_{2C_r} de las dos nuevas nubes.
Sustituir el prototipo de C_r por los dos nuevos prototipos, p_{1C_r} y p_{2C_r}: $P \leftarrow P - \{p_{C_i}\} \cup \{p_{1C_r}, p_{2C_r}\}$
Reclasificar todos los vectores en N_{norm}^{prot} de acuerdo a un criterio de mínima distancia y a los prototipos, P, establecidos.
Obtener la distorsión media, D_m: $D_m = \frac{1}{N_p} \sum_{i=1}^{N_p} \frac{1}{N_{C_i}} \sum_{\forall q \in C_i} d(q, p_{C_i})$
Finalización: Se usa el conjunto de prototipos P para representar el espacio N_{norm}^{prot}.

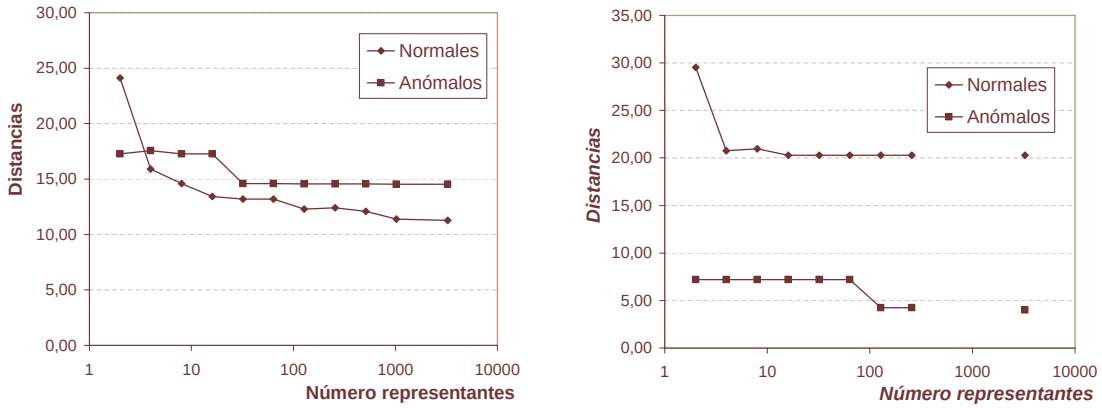


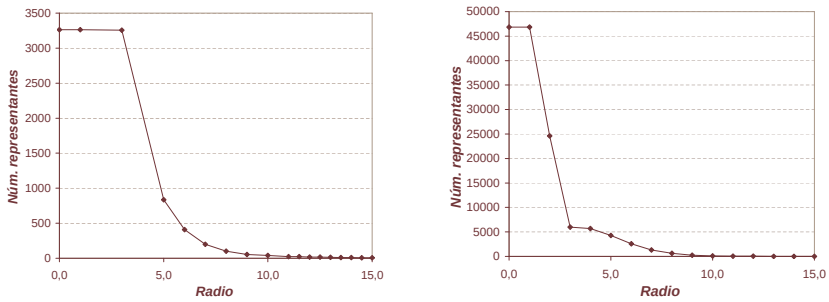
Figura 4: Evolución de los índices de anormalidad de los conjuntos de evaluación con el número de prototipos considerado para el algoritmo N1 para HTTP (izda.) y DNS (dcha.).

dificaciones en el algoritmo *k-medias*: la división de la nube se realiza tomando los puntos más distantes de la nube como centroides perturbados y, una vez obtenidos todos los representantes de todas las nubes, se reclasifican todos los vectores existentes en función de un criterio de mínima distancia (Fig. 3). Este último paso no es necesario en el algoritmo *k-medias* ya que, por construcción, está garantizado que los vectores se encuentran más próximos al centroide de su nube que al de las restantes. Sin embargo, como se ha apuntado con anterioridad, esto no es cierto en el caso que nos ocupa. El algoritmo resultante, denominado N1, se detalla en la Tabla 2.

Como se observa en la Fig. 2.b, la convergencia es más suave que en el caso del algoritmo *k-medias* modificado, aunque sigue sin garantizarse la monotonicidad de la misma.

Por otra parte, los resultados experimentales obtenidos para HTTP y DNS (Fig. 4) muestran un buen comportamiento de este algoritmo, consiguiéndose una reducción significativa del número de representantes sin degradación del rendimiento del sistema. Como se observa en la Fig. 4, la sucesiva reducción en el número de representantes (de derecha a izquierda en la gráfica), mantiene la separación entre los índices de anormalidad del tráfico normal y el de ataque para reducciones considerables en el tamaño del modelo.

Sin embargo, este algoritmo puede resultar muy costoso, ya que implica un considerable número de cálculos de distancias en cada iteración o, alternativamente, el precálculo de todas ellas. Este problema se agrava si el número de elementos en el modelo es elevado. En el caso del proto-



a) b)

Figura 5: Evolución del algoritmo $N2$: número de elementos en el modelo en función del radio para: a) HTTP y b) DNS.

colo DNS, con el modelo obtenido han aparecido problemas graves de capacidad y gestión en los equipos disponibles debido, fundamentalmente, a limitaciones de memoria.

4.3. Algoritmo de agrupamiento N2

Las limitaciones que presenta el algoritmo de agrupamiento N1 nos llevan a proponer una solución ad hoc basada en una idea muy simple y que, simultáneamente, presenta una complejidad de implementación muy reducida.

El algoritmo, denominado N2, se basa en la utilización de un parámetro de entrada, R, denominado *radio*. La idea básica consiste en seleccionar un elemento del modelo y eliminar todos los elementos del mismo situados a una distancia inferior a dicho radio, asumiendo que el elemento original representa adecuadamente a los elementos suprimidos. La iteración del procedimiento descrito de forma adecuada originará una reducción en el tamaño del modelo. La operación se puede describir mediante el siguiente pseudocódigo:

```
Definiciones:
    R: radio
    N_Mod: modelo original
    N_pMod: modelo reducido
Inicialización: N_pMod = N_Mod
Para cada carga útil p en N_pMod
    Para cada carga útil q ≠ p en N_pMod
        Si d(p,q) ≤ R
            Eliminar q de N_pMod
```

Evidentemente, el tamaño final del modelo es función del radio elegido. A diferencia del algoritmo N1, en el que se puede fijar el número de elementos que se desea conservar en el modelo (como criterio alternativo a un umbral en la distorsión media), en este caso no se conoce ni se puede fijar a priori la reducción.

La Fig. 5.a muestra los resultados de reducción obtenidos para el modelo HTTP en función del radio. Para valores de R inferiores a 3 se observa que no se produce reducción en el tamaño del modelo. A partir de $R = 5$ se producen reducciones considerables sin pérdida significativa de rendimiento (Fig. 6.a), ya que se mantiene la separación entre el tráfico normal y el anómalo. Sin embargo, si se intenta conseguir reducciones drásticas se observa que se produce solapamiento entre los rangos del índice de anormalidad del tráfico normal y el anómalo, lo que genera errores de detección y falsas alarmas. Los resultados para DNS son análogos, como se muestra en las Figs. 6.b y 6.b.

4.4. Evaluación

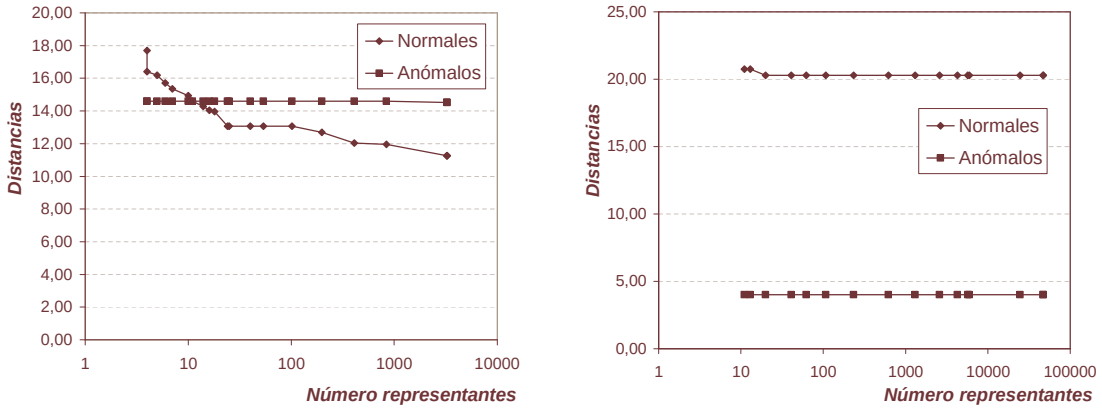
Los resultados obtenidos mediante ambos algoritmos pueden proporcionar una importante mejora en el rendimiento sin reducir la eficacia de la detección. A modo de ejemplo, considerando los modelos reducidos para HTTP con 25 prototipos, que aún proporcionan resultados de detección satisfactorios, y utilizando los tiempos de cómputo indicados en la Sección 2, se consigue una reducción del tiempo de detección del valor original $t_{detec} = 15{,}76ms$ a un valor de $t_{detec} = 0{,}121ms$. Con estos valores, el detector podría gestionar 8264 peticiones por segundo, obteniéndose un incremento significativo sobre el modelo original.

En el caso de DNS la reducción es todavía mayor. Con $N = 107$ prototipos, el sistema puede gestionar 1937 peticiones por segundo, frente a las 5 del modelo original.

Como se ha mostrado, ambos algoritmos presentan un buen comportamiento, si bien presentan ventajas e inconvenientes relativos. El algoritmo N1, como se ha mencionado, presenta un elevado coste computacional, al contrario que el algoritmo N2, que resulta bastante liviano. Ambos pueden proporcionar resultados subóptimos, debido a la naturaleza del problema y la solución adoptada. Por otra parte, N2, en su formulación actual, adolece de cierto grado de aleatoriedad, ya que el orden en el que se procesan los prototipos afecta al resultado final.

5. Conclusiones

Se han propuesto y evaluado dos algoritmos de agrupamiento que han mostrado que cumplen el doble objetivo planteado. En primer lugar, reducen de forma efectiva el tamaño de los modelos de normalidad usados por el detector N3 para



a) b)

Figura 6: Evolución de los índices de anormalidad de los conjuntos de evaluación con el algoritmo N2: a) Para HTTP y b) para DNS.

los protocolos HTTP y DNS, lo que mejora el rendimiento computacional del sistema y facilita su implantación en entornos reales. En segundo lugar, ambos procedimientos han mostrado que la reducción puede ser significativa sin degradar la eficacia del detector.

Sin embargo, este trabajo debe entenderse como una primera aproximación al problema, debiendo abordarse aspectos como la sistematización de los mecanismos de reducción, en el sentido de que se proporcionen métodos eficaces que sean capaces de determinar por sí mismos el grado de reducción idóneo, o la utilización de procedimientos de agrupamiento alternativos, potencialmente más eficaces.

Agradecimientos

Este trabajo ha sido parcialmente subvencionado por el MECD a través del PNFPU (referencia AP2001-3805) y el MCYT a través del proyecto SERVIRA (TIC2002-02798, 70 % fondos FEDER).

Referencias

[1] Denning, D., *An Intrusion-Detection Model*, in IEEE Transactions on Software Engineering, Vol. SE-13, No. 2, pp. 222-232, February, 1987.

[2] McHugh, J.; *Intrusion and Intrusion Detection*, Internation Journal on Information Security, Vol. 1., No. 1., pp.14-35, 2001.

[3] Axelsson, S.; *Intrusion Detection Systems: A Survey and Taxonomy.* Technical Report 99-15, Department of Computer Engineering, Chalmers University of Technology, Goteborg.

[4] Allen, J. y otros; "State of the Practice of Intrusion Detection Technologies". Technical Report CMU/SEI-99-TR-028, Software Engineering Institute, Carnegie Mellon Univ., 2000.

[5] Estevez-Tapiador, J. M. y otros; *Stochastic Protocol Modeling for Anomaly-Based Network Intrusion Detection*, Proc. 1st IEEE International Workshop on Information Assurance (IWIA'03), pp. 3-12, Darmstadt, 2003.

[6] Krügel, C., Toth, T., and Kirda, E.; *Service Specific Anomaly Detection for Network Intrusion Detection*, Proc. 17th ACM Symp. on Applied Computing (SAC), pp. 201-208, Madrid, Spain, 2002.

[7] J. M. Estévez Tapiador; *Detección de intrusiones en redes basada en anomalías mediante técnicas de modelado de protocolos*, Tesis doctoral, Universidad de Granada, 2004.

[8] J. M. Estévez-Tapiador y otros; *N3: A geometrical approach for network intrusion detection at the application layer*, ICCSA 2004, LNCS 3043, pp. 841-850, 2004.

[9] J. M. Estévez Tapiador y otros; *Detección geométrica basada en anomalías de ataques sobre HTTP*, en *Avances en Criptología y Seguridad de la Información*, Ed. Díaz de Santos, pp. 455-464, 2004.

[10] Gusfield, D.; *Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology*, ISBN: 0521585198, Cambridge University Press, 1997.

[11] Athanasiades, N. y otros; *Intrusion Detection Testing and Benchmarking Methodologies*, Proc. 1st IEEE International Workshop on Information Assurance (IWIA'03), Darmstadt (Germany), 2003, pp. 63-72.

[12] Lippmann, R. y otros; *The 1999 DARPA Offline Intrusion Detection Evaluation*, Computer Networks, Vol. 34, No. 4, pp. 579-595, 2000.

[13] McHugh, J.; *Testing Intrusion Detection Systems: A Critique to the 1998 and 1999 DARPA Intrusion Detection Evaluations as Performed by Lincoln Laboratory*, ACM Transactions on Information and Systems Security, Vol. 3. No. 4, pp. 262-294, 2000.

[14] *Arachnids: Advanced Reference Archive of Current Heuristics for Network Intrusion Detection Systems.* http://www.whitehats.com/ids, 2003.

[15] Linde, A., Buzo, Y., Gray, R.M., *An Algorithm for Vector Quantizer Design.* IEEE Trans. on Communications, COM-28(1), pp. 84-95, 1980.

Análisis de mecanismos software para la captura de tráfico en red

Igor Delgado[1], Armando Ferro[2], Alfredo Beaumont[3], Alex Muñoz[4]
Networking, Quality and Security Research Group
Departamento de Electrónica y Telecomunicaciones. Universidad del País Vasco
ETSI de Bilbao C/Alameda Urquijo, s/n C.P.:48013 – Bilbao
Teléfono: 94 601 73 08. Fax: 94 601 42 59
E-mail: {jtbdegai[1]|jtbbesaa[3]}@aintel.bi.ehu.es, {jtpfevaa[2]|jtpmumaa[4]}@bi.ehu.es

Abstract. *Analysis of network traffic has become a really important task for many fields, from IDS to QoS evaluation. The great increase in the speed of networks, much faster than processor's performance increase, is making difficult to handle the processing of all network packets. One way to solve this problem is to move the processing to the hardware elements. This is a very expensive solution and, in some cases, improving the performance of the capturing software may be enough. Although current general purpose operating systems have been improving the performance of network processing, there is still place for improvement. In this paper we present in detail the way followed by a packet through the operating system until it is processed at user level. Some empirical measurements has been made along the packet journey in order to find the bottlenecks of the capturing process as well as discussing some ways to further increase the performance of the capturing process.*

1 Introducción

Los sistemas de captura de tráfico de red juegan un papel fundamental en multitud de campos de la telemática como la detección de intrusión, la medida de parámetros de calidad de servicio, la ingeniería de tráfico o la detección de virus. En todas estas disciplinas se emplean sistemas que capturan y analizan el tráfico de red para obtener estadísticas. Si el procesamiento que se le aplica a cada paquete es computacionalmente costoso, el sistema no será capaz de procesar todos los paquetes de la red produciéndose pérdidas. En los últimos años la evolución de las infraestructuras de red ha sido más rápida que la del hardware de captura con lo cual ese problema se ha visto fuertemente acrecentado.

Por esta razón, a los sistemas de captura de tráfico se les exige un alto rendimiento evitando en lo posible el consumo innecesario de los recursos disponibles. La mayor parte de los sistemas de análisis de tráfico basados en software se suelen construir en el nivel de usuario al ser más fácil su desarrollo y portabilidad. El software de captura de paquetes suele basarse en las facilidades proporcionadas por el sistema operativo. Éste es el encargado de capturar los paquetes y transportarlos desde la interfaz de red hasta el área de memoria de usuario. En un sistema operativo (SO, en adelante) convencional [1] el mecanismo de captura de datos no suele estar optimizado. Esto es lógico ya que la mayoría de los SOs son de propósito general y no están específicamente diseñados para realizar una captura eficiente de tráfico. Si se piensa en diseñar un buen sistema de captura de tráfico a partir de un SO de propósito general, es importante conocer cómo funciona y saber cómo mejorar su diseño para alcanzar un rendimiento óptimo.

En la fase de captura de tráfico el paso de los paquetes desde la tarjeta de red hasta el nivel de usuario supone un coste computacional muy importante. Una de las líneas de investigación en este campo que más relevancia tienen hasta ahora está centrada en el desarrollo de tarjetas de captura de datos con hardware especial [2] [3]. El hardware está específicamente diseñado para la captura de tráfico permitiendo alcanzar velocidades de hasta 10 Gbps. De esta forma, es posible reservar la capacidad de la CPU para aplicar algoritmos de tratamiento software complejos a los paquetes. Sin embargo, estas tarjetas tienen el inconveniente de su alto costo.

En muchos casos donde las exigencias de captura no son tan extremas es posible que una optimización del software de captura sea suficiente para analizar todo el tráfico de red. Esto permite utilizar plataformas hardware de propósito general como sondas para captura de tráfico.

En este trabajo se presenta un estudio de las limitaciones software de un sistema de captura de tráfico convencional. En las secciones 2 y 3 se estudian en detalle las etapas de la captura pasiva de tráfico. A continuación, en la sección 4, se presentan datos recogidos del análisis realizado sobre un prototipo experimental de laboratorio. Ello permite estimar el coste computacional de las diferentes etapas de procesamiento y su influencia en el diseño adecuado de un sensor para análisis de tráfico. A partir de esa información se realizan consideraciones de interés que permitirán determinar cuáles son los puntos más susceptibles de mejora en el diseño de un sistema software de captura de tráfico. Finalmente, en la sección 5 las conclusiones más relevantes de nuestro trabajo son expuestas.

2 Mecanismos de captura de tráfico

El creciente aumento del ancho de banda de las redes locales ha exigido a los sistemas operativos una evolución en el diseño software de sus sistemas de red. Si hace años las redes locales intercambiaban información a pocos Mbps ahora fácilmente llegan a los Gbps. A lo largo de este camino se han ido desarrollando diferentes soluciones de propósito general que se describirán brevemente en esta sección.

2.1 Mecanismos de captura basados en interrupciones

Los primeros sistemas operativos utilizaban un mecanismo basado en interrupciones para planificar las operaciones de red. Es decir, cada vez que llegaba un paquete a la tarjeta de red, ésta generaba una interrupción y se ejecutaba la rutina de atención correspondiente a esa interrupción. En esa rutina el kernel realizaba, sobre el paquete recién capturado, todas las operaciones de red que sean necesarias. Una vez que la rutina había finalizado se continuaba con la tarea interrumpida. Este proceso se repetía con cada paquete que llega a la tarjeta.

Las ventajas de este mecanismo son la facilidad en su implementación y la baja latencia con la que los paquetes capturados se entregan al nivel de usuario. Sin embargo, tiene un inconveniente muy importante. Cuando la frecuencia de llegada de paquetes al sistema es alta, la mayor parte de los recursos de la CPU se consumen en las rutinas de atención a las interrupciones provocadas con la llegada de esos paquetes. El resto del tiempo disponible se repartirá entre las tareas activas del sistema. Si la tasa de llegada crece mucho, las interrupciones monopolizarán el uso de la CPU impidiendo la ejecución de cualquier otra tarea y llegando a lo que en [4] se denominó un *livelock*. De esta situación no se puede salir hasta que se reduzca la tasa de llegada de paquetes.

2.2 Mecanismos de captura basados en *polling*

Para evitar los *livelocks* se pueden emplear mecanismos de *polling* [10], es decir, el SO pregunta periódicamente a la tarjeta si ha llegado algún paquete. En caso afirmativo se procesarían todos los paquetes que hayan llegado hasta ese momento. De esta forma aunque la tasa de llegada sea muy elevada el sistema no se bloquea ya que es el propio kernel quien controla el consumo computacional del dispositivo de red y, por tanto, el número de paquetes que entran al sistema.

Este mecanismo también tiene inconvenientes como el retardo con la que los paquetes son atendidos por el sistema operativo. Además, cuando la tasa de llegada es elevada, durante cada *polling* el SO captura varios paquetes consumiendo recursos por cada uno de ellos. Dado que la tasa es alta, es probable que el sistema que vaya a analizar esos paquetes no sea capaz de gestionar todos ellos debido a la escasez de ciclos de CPU. Esto lógicamente es ineficiente ya que habría paquetes que han consumido recursos de forma innecesaria al no llegar nunca a un sistema de análisis de los paquetes a nivel de usuario.

2.3 Mecanismos de captura mixtos

Dadas las ventajas y los inconvenientes de los mecanismos anteriores J. Mogul [4] presentó un mecanismo mixto capaz de evitar los *livelock* para los sistemas BSD 4.2 y a la vez reducir los consumos innecesarios debidos al *polling*. Para ello utilizó la técnica denominada coalescencia de interrupciones en la que el kernel procesa varios paquetes por cada interrupción.

Cuando un paquete llega al sistema se produce una interrupción hardware. Como siempre tras la interrupción se llamará a la rutina de atención correspondiente en la que se indicará al planificador que planifique una tarea de captura de paquetes y se deshabilitarán las interrupciones.

Cuando el sistema lo considere oportuno planificará una tarea del kernel encargada de gestionar los paquetes recibidos y hacer un *polling* sobre la tarjeta de red. El número de paquetes capturado variará en función del tiempo que haya pasado desde que se produjo la interrupción hasta que el planificador haya podido ejecutar esta tarea. De esta forma se capturan varios paquetes con el coste de una sola interrupción mejorando el rendimiento. Una vez finalizada esta tarea se vuelven a habilitar las interrupciones y el procedimiento se inicia de nuevo.

Este mecanismo es más eficiente ya que cuando la frecuencia de llegadas sea baja se comportará prácticamente como un sistema basado en interrupciones y cuando aumente se capturarán más paquetes por interrupción.

El mayor inconveniente de este mecanismo es la latencia que se produce desde que el paquete llega al sistema hasta que finalmente es atendido por la tarea del kernel encargada de su procesamiento.

3 Análisis del sistema de captura en Linux

En esta sección se explicará detalladamente cuál es el recorrido de un paquete desde que es capturado por una tarjeta de red Ethernet hasta que llega al nivel de usuario para ser analizado por el proceso correspondiente. Para un sistema de captura de datos sólo tiene relevancia el procedimiento de recepción de información del sistema de red por lo que no se va a considerar el procedimiento de emisión. Los paquetes capturados en modo promiscuo no atraviesan la pila TCP/IP ya que no pertenecen a

ninguna conexión establecida con el sistema de captura por lo que tampoco se tendrá en cuenta. Se ha elegido el kernel de Linux como base para el estudio debido a su carácter abierto y que es utilizado en muchos sistemas de captura. A partir del kernel 2.4.20, la implementación de los mecanismos de captura ha sido modificada para que sea más uniforme y eficiente. Con la inclusión de la *New API* (NAPI) [5] los sistemas Linux pasaron a implementar un mecanismo de captura mixto que es el que se explica a continuación.

La captura se ha divido en dos partes atendiendo al hilo o proceso que realiza el procesamiento (ver Figura 1). En la primera de ellas, denominada tratamiento de kernel, se engloban todas las tareas que realiza el kernel cada vez que captura un paquete. El tratamiento de kernel se corresponde con los dos cuadros de la izquierda en la Figura 2.

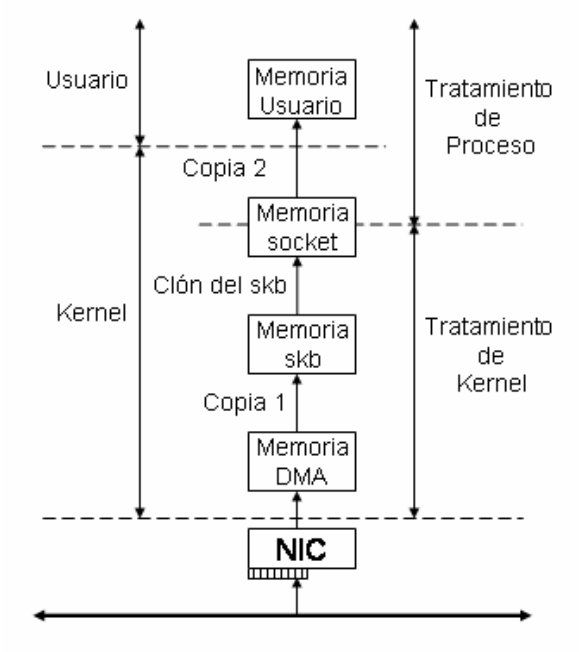


Figura 1: Captura en Linux (kernel >=2.4.20)

En la segunda parte de la captura, denominada tratamiento de proceso, se comentarán las acciones llevadas a cabo por el proceso de usuario para recibir los paquetes capturados por el sistema. El tratamiento de proceso se corresponde con los dos cuadros de la derecha en la Figura 2.

En los siguientes párrafos se hablará de acciones, representadas mediante cuadros en la figura 2, y de etapas, representadas mediante círculos, también en la Figura 2. Las etapas son transiciones del paquete capturado a través del sistema. Por otro lado, las acciones se corresponden con todas operaciones necesarias para que el mecanismo de captura funcione correctamente.

3.1 Tratamiento de kernel

Una vez que la tarjeta de red ha sido inicializada correctamente, empieza a analizar las señales que recibe del medio. Tras detectar el preámbulo de una trama comienza su almacenamiento en una zona de memoria interna propia de la tarjeta. Después de que lleguen todos los bits de la trama, la NIC calcula el CRC del nivel Ethernet. Si no es correcto el paquete es desechado, en caso contrario, cuando el bus I/O[1] esté libre, el paquete es transferido a la memoria principal del sistema mediante transferencias DMA (etapa A). El tamaño del buffer de la memoria principal en el que el DMA deposita los paquetes de red depende del hardware de la tarjeta y del propio dispositivo DMA Una vez finalizada la transferencia se genera automáticamente una interrupción hardware, *hardirq*. Hasta este momento, todas las acciones se han llevado a cabo sin que el kernel tenga constancia de ello y sin consumo de CPU.

Cuando el kernel detecta la interrupción, se ejecuta la rutina de atención a la interrupción del driver (acción 1). En esta rutina se realizan tres acciones: deshabilitar las *hardirq*, planificar la ejecución de una *softirq* que procese el paquete, y, por último, confirmar a la tarjeta la correcta recepción de la interrupción. La rutina de atención debe ser lo más corta posible para no ralentizar en exceso la tarea que se estuviera ejecutando en la CPU a la hora de la interrupción. En la *hardirq* no se realiza acción alguna sobre los paquetes.

Realmente, la primera acción sobre los paquetes capturados se producirá cuando la *softirq* sea invocada (acciones 2, 3 y 4). El objetivo de la *softirq* es trasladar los paquetes desde la zona de memoria donde el DMA ha depositado los paquetes hasta la cola del socket que esté esperando los datos. Para ello, tras reservar el espacio necesario, en área de kernel, copia el contenido del paquete desde el buffer DMA (etapa B). Ésta es la primera copia del contenido del paquete realizada por la CPU. Seguidamente, se indica al controlador de DMA que ese paquete ya ha sido copiado y que el espacio del buffer DMA utilizado puede ser sobrescrito.

A continuación, aún dentro de la misma *softirq* se encola una referencia del paquete en la cola del socket (etapa C). Es importante aclarar que no se copia todo el paquete sino únicamente información asociada al paquete y una referencia a su posición (clon). Un efecto perjudicial observado en las pruebas realizadas es que cuando la cola del socket se llena, se producen pérdidas de paquetes. Luego, cuanto más grande sea esa cola mayor será la resistencia del sistema de captura.

[1] Generalmente buses PCI o similar.

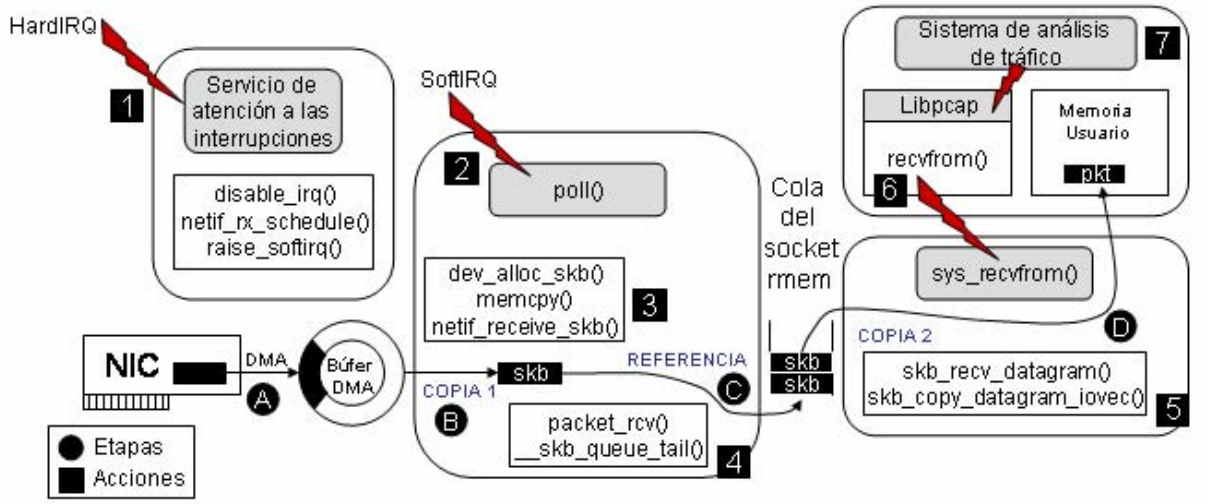


Figura 2: Diagrama general del mecanismo de captura en Linux

Sin embargo, como se comentará en apartados posteriores esta cola es útil sólo ante las ráfagas de tráfico puntuales siendo inoperante cuando el tráfico es sostenido.

Una vez guardada la referencia al paquete en la cola del socket, la *softirq* volverá a por otro paquete y realizará las mismas acciones descritas hasta ahora (etapas B y C, acciones 3 y 4). Esto se repetirá hasta que no haya más paquetes en el área de DMA utilizada por la tarjeta de red o hasta que complete un cupo máximo establecido (en general 64 paquetes). Por tanto, cuando la tasa de llegadas sea muy alta, la *softirq* siempre tendrá un paquete disponible para tratar por lo que normalmente finalizará tras completar su cupo máximo de paquetes.

Por último, antes de salir de la *softirq* es necesario volver a habilitar las *hardirq* para que se pueda producir de nuevo el procedimiento de captura.

3.2 Tratamiento de proceso

Todas las acciones que se describen en esta sección se realizan en el contexto de un proceso pero no se ejecutan exclusivamente en el entorno de nivel de usuario, sino que parte de ellas se ejecutan en el entorno de nivel de kernel (ver Figura 2). Esta última parte se corresponde con la actividad que se desarrolla en el kernel al atender las demandas específicas de un proceso de usuario invocadas a través de llamadas al sistema en la interfaz del socket.

La cola del socket es el punto de unión entre el tratamiento de kernel y el tratamiento de proceso. En un sistema de análisis de tráfico, el proceso de usuario es el que tiene que tomar la iniciativa y hacer una petición de los paquetes de red. Para ello se abre un socket en modo promiscuo y se ejecuta `recvfrom` (acciones 6 y 7). Con ello se dispara la llamada al sistema `sys_recvfrom` (acción 5). Esta función comprueba si hay algún paquete en la cola del socket quedándose dormido a la espera en caso contrario. Cuando el SO haya recibido un paquete y lo haya encolado siguiendo el procedimiento descrito en la sección anterior, el proceso de usuario será despertado. A continuación, desencolará la referencia al paquete con su información asociada (clon) y se copiará el contenido del paquete de la memoria del kernel a la memoria de usuario (etapa D). Ésta es la segunda copia del contenido del paquete durante su recorrido.

Existen muchas aplicaciones que en lugar de configurar el socket directamente hacen uso de Libpcap [6]. Libpcap es una librería que ofrece una interfaz común para la captura de paquetes en diferentes SOs que proporciona facilidades para configurar filtros en el kernel y da estadísticas sobre el proceso de captura. Sin embargo, el uso de Libpcap como solución de diseño para un sistema de captura de tráfico introduce consumos computacionales que limitan su rendimiento.

Por tanto, en la etapa A se trasladan los paquetes de la memoria de la tarjeta de red a la zona de memoria DMA. En la etapa B, se copia el paquete de la zona DMA a otra zona de memoria del kernel y se crea el *socket buffer* (skb). A continuación, en la etapa C se encola una referencia a ese skb en la cola del socket. Por último, en la etapa D se copia el paquete desde la cola del socket a la memoria de usuario.

4 Evaluación del sistema de captura

Una vez descrito el funcionamiento de un sistema de análisis de tráfico, en esta sección se evaluará el consumo en cada una de las etapas y el tamaño de los bufferes por la que pasa un paquete en la captura. De esta forma será posible determinar cuáles son los cuellos de botella y los posibles puntos de mejora en la captura de paquetes.

4.1 Entorno de pruebas

Los resultados que se muestran en los siguientes apartados se corresponden con unas pruebas realizadas en una red local Gigabit Ethernet aislada del exterior. Se ha utilizado una máquina dual AMD Opteron a 1,8 GHz, con 2048 MB de RAM y una tarjeta Gigabit Broadcom como inyector de tráfico. La máquina receptora tiene un procesador Intel Xeon a 2,4 GHz con 1024 MB de RAM y una tarjeta Broadcom. Ambas máquinas se encuentran conectadas a un switch Gigabit. En las pruebas realizadas se pretende medir el rendimiento de las máquinas en casos extremos, por lo que se inyectan paquetes pequeños de 60 bytes hasta alcanzar velocidades de 450.000 paquetes por segundo (lo máximo que llega a dar la máquina inyectora, no se han utilizado más máquinas porque 450.000 pps es suficiente para el estudio llevado a cabo en el artículo).

Los sistemas operativos empleados son Debian GNU/Linux con kernels 2.6.10. El driver de las tarjetas Broadcom es el Tigon3. Para calcular los tiempos de cada etapa se ha utilizado la instrucción `rdtscl` que da el número de ciclos de reloj de la CPU. Este tiempo tiene una precisión muy elevada dependiente de la frecuencia de la CPU del sistema. Cada unidad de este registro de 64 bits se corresponde con la inversa de la frecuencia de reloj de la CPU [9], es decir, para una CPU de 2,4 GHz cada ciclo dura unos 0,42 nanosegundos. Los tiempos se expondrán tanto en ciclos como tiempo ya que de esta manera los cálculos pueden ser extrapolados a otras máquinas con procesadores a diferente frecuencia.

4.2 Análisis del tamaño de los bufferes

Primeramente se realizará un análisis del tamaño de los bufferes que atraviesa cada paquete y su influencia en la tasa de pérdidas en el análisis.

4.2.1 Buffer DMA y memoria de kernel

El primer buffer a estudiar es el área de DMA. Todos los paquetes son transferidos desde la memoria interna de la tarjeta hasta la zona correspondiente del DMA emplazada en la memoria del sistema sin consumo de CPU mediante el motor de transferencia DMA. El kernel reserva espacio para copiar la información asociada al paquete que acaba de llegar vía DMA. Esta información suele consistir en una copia del contenido, marca del tiempo de llegada, información del dispositivo al que está asociado, etc. En un principio el límite en la reserva es el impuesto por la memoria física del sistema, es decir, el kernel no tiene un espacio limitado para almacenar los paquetes que llegan sino que los irá albergando de forma dinámica hasta que no disponga de más memoria. Las pruebas realizadas reflejan claramente que, hasta las velocidades a las que se han inyectado los paquetes (unos 450.000 pps), no se llega a llenar nunca la memoria del sistema por lo que éste no es un factor limitador en la captura.

El tamaño del área de DMA no suele ser excesivamente grande y es dependiente del sistema operativo y del fabricante. Generalmente el motor DMA es capaz de extraer todos los paquetes de la tarjeta de red, y es el kernel el que no es lo suficientemente rápido como para recoger todos los paquetes del buffer. En ese caso el buffer se completará y se sobrepasará produciéndose pérdidas en la captura. La siguiente tabla muestra este efecto.

Tabla 1: Porcentaje de pérdidas motor DMA.

Porcentaje de pérdidas en el DMA					
	50	100	200	300	400
Sin carga	0	0	0	0	0
Sin análisis	0	0	23,2	40,2	57,5
Con análisis	0	37,17	39,1	50,5	62,3

Las columnas representan la velocidad de inyección en miles de paquetes por segundo. Las filas, el porcentaje de pérdidas que se produce en la transición del buffer DMA a la memoria del kernel a la tasa de llegada correspondiente. La primera fila se corresponde con el porcentaje de pérdidas cuando el sistema está libre de carga, es decir, sólo hay tratamiento de captura en el kernel, sección 3.2. En la segunda fila, se indican las pérdidas de paquetes en el supuesto de que el proceso de usuario los captura pero no realiza ningún análisis (tratamiento de kernel más tratamiento de proceso, pero sin análisis de los paquetes, sección 3.2). Y, por último, en la tercera fila se indican las pérdidas de paquetes supuesta una carga de análisis adicional aplicada a cada paquete durante el tratamiento de proceso.

A medida que la carga del sistema aumenta, el porcentaje de paquetes perdidos es mayor. Esto puede deberse a varios factores. El primero de ellos es el bus que conecta la tarjeta con la memoria principal del sistema (RAM). En las pruebas realizadas, este bus es PCI-X y tiene un ancho de banda de 4,2 GBps [11], suficientemente grande como para no provocar pérdidas a las velocidades de inyección. Además, el bus no está compartido. Otro elemento que podría ser el causante de las pérdidas es la memoria RAM. Sin embargo, la memoria utilizada tiene un ancho de banda entorno a los 2-3 GBps por lo que tampoco supone un cuello de botella. El último factor por considerar es el procesador (CPU). Cuando las velocidades son bajas, 50.000 y 100.000 pps no se producen pérdidas porque la CPU es suficientemente potente como para poder dedicar los ciclos necesarios para el tratamiento de kernel y el de proceso de todos los paquetes. En cuanto las velocidades de inyección aumentan la CPU empieza a saturarse y surgen las pérdidas debido a que cada vez le quedan menos ciclos libres al sistema para recoger los paquetes del buffer DMA. Por tanto, mientras el SO tenga ciclos libres es capaz de sacar todos los paquetes de la memoria del DMA.

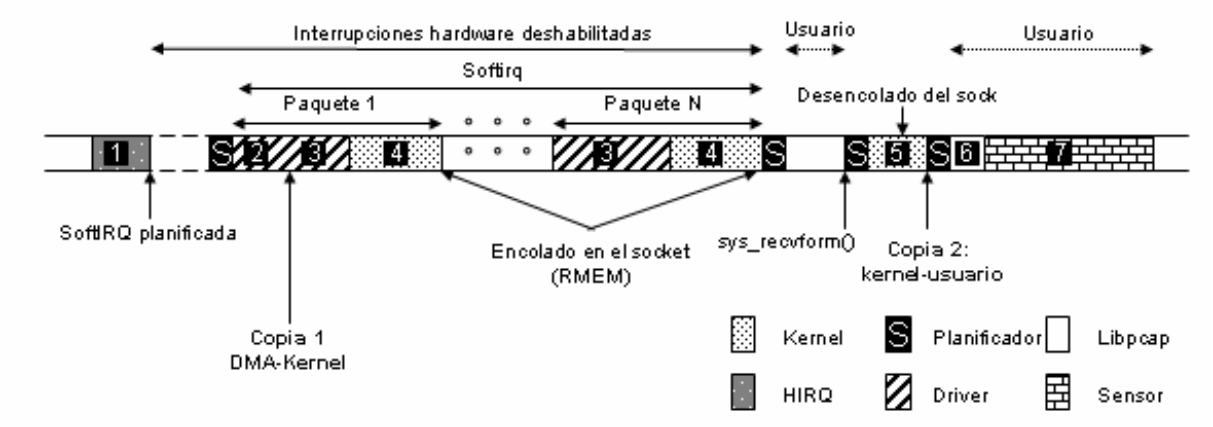


Figura 3: Evolución temporal del mecanismo de captura para Linux

Esto pasa cuando sólo se realiza un tratamiento de kernel sobre los paquetes, al ser menos costoso que el tratamiento de kernel y de proceso juntos, el sistema es capaz de capturar todos los paquetes.

Estas medidas se han realizado modificando el código del driver y del kernel para calcular las estadísticas, tanto de paquetes perdidos como de tiempos (con la instrucción `rdtscl`).

4.2.2 Buffer del socket

El siguiente buffer es la cola del socket que espera los paquetes. En Linux se puede ver el tamaño máximo del buffer del socket donde se almacenan los paquetes `/proc/sys/net/core/rmem_default`. En este buffer no se copia todo el paquete porque sería necesario realizar otra copia del contenido entre dos zonas de memoria del kernel con el consumo innecesario que eso conlleva. En lugar de eso, sólo se encola una referencia al paquete junto con "metainformación" asociada al mismo. Esto tiene una longitud final de unos 270 bytes.

Tabla 2: Porcentaje de pérdidas tras inyectar 15 millones de paquetes.

rmem ↓	100	200	300	400
0,15 MB	0	96,79	96,58	96,51
1,5 MB	0	87,87	96,56	96,64
15 MB	0	86,35	95,66	95,54

La Tabla 2 representa el porcentaje de pérdidas en el sistema cuando se inyectan 15.000.000 paquetes de forma consecutiva a las velocidades que van desde 100.000 pps hasta 400.000 pps.

Tabla 3: Porcentaje de pérdidas tras inyectar ráfagas de 50.000 paquetes.

rmem ↓	100	200	300	400
0,15 MB	0,16	95,78	95,35	93,59
1,5 MB	0	71,58	73,73	74,14
15 MB	0	0	0	0

Como se puede observar, un aumento en la longitud de la cola del socket (desde 150.000 hasta 15 millones de bytes) no mejora el número de paquetes capturados para velocidades sostenidas. Por el contrario, si se inyectan ráfagas cortas y la longitud de la cola del socket (rmem) es suficientemente grande, el porcentaje de pérdidas disminuye hasta reducirse a 0. En estas pruebas las ráfagas son de 50.000 paquetes.

Por tanto, la principal función de esta cola es hacer frente a las ráfagas instantáneas tan comunes en el tráfico en Internet. Cuando es posible alojar todos los paquetes de una ráfaga en la cola del socket el sistema es capaz de hacerse cargo de todo el tráfico y analizar todos los paquetes sin perder ninguno. Cada clon del socket necesita 270 bytes, el producto de 50.000 por 270 es menor que 15.000.000 por lo que no se pierde ningún paquete. En cambio cuando la cola del socket se llena el porcentaje de pérdidas es muy grande. Un factor de diseño a considerar sería la reducción del tamaño de esas referencias (clones, desde los 270 bytes a un valor menor) pues con ello se podría disponer de más capacidad de almacenamiento.

Cuando la tasa de llegadas es baja, 100.000 paquetes por segundo, el sistema también es capaz de analizar todos los paquetes. Esto se debe a que el proceso de usuario es suficientemente rápido como para desalojar la cola del socket a la misma velocidad a la que se llena. En cuanto la relación entre estas velocidades se descompensa surgen las pérdidas.

4.3 Tiempos de captura

La línea temporal de una captura genérica se muestra en la Figura 3. Cuando llega un paquete, la tarjeta produce una interrupción hardware. Esta interrupción tiene una duración bastante estable de 640 ciclos de reloj que se corresponde en nuestra máquina con 0,27 microsegundos.

Una vez terminada la interrupción, el procesamiento del paquete continúa cuando el planificador ejecuta la *softirq* correspondiente (etapas B y C). En una *softirq* la duración media de cada captura es de 5.715 ciclos por paquete, 2,37 microsegundos. También se han

calculado tiempos intermedios durante la ejecución de la misma. Los más relevantes son: el tiempo medio que consume el copiado desde el buffer DMA a la memoria RAM (etapa B), que tiene un valor de 2.900 ciclos por paquete, 1,22 microsegundos, y, el tiempo medio utilizado durante la creación y encolado del clon del paquete (etapa C), con un valor de 1.550 ciclos por clon, 0,64 microsegundos.

El resto de tiempo hasta completar los 2,37 microsegundos (0,51 microsegundos) se emplea en tareas variadas de bajo consumo como la actualización de contadores o llamadas a funciones. Hay que destacar que este valor es por cada paquete y que por cada *softirq* normalmente se captura más de un paquete. Cuando la tasa de llegadas es alta en cada *softirqs* se captura más de un paquete por lo que el tiempo medio por *softirq* es mayor.

Una vez encolado el paquete, el proceso de usuario es el encargado de desencolar los paquetes y procesarlos. Para ello habrá realizado una llamada a `recvfrom()` y se habrá quedado dormido a la espera de un paquete. La copia de cada paquete desde el kernel al nivel de usuario (etapa D) depende ligeramente de la longitud del paquete, cuanto mayor sea el paquete más costará copiarlo. La siguiente tabla muestra en la primera fila diferentes tamaños de paquete en bytes. En la siguiente se muestran el número de ciclos que se emplea en hacer una copia y la tercera es el mismo valor en microsegundos.

Tabla 4: Efecto de la longitud del paquete en la copia del kernel al usuario (etapa D).

Longitud de los paquetes		
60 bytes	600 bytes	1500 bytes
4.790 ciclos	4.860 ciclos	5.090 cilcos
2,01 µs	2,04 µs	2,14 µs

Esta copia puede ser eliminada usando `mmap` [7] ya que evita la copia de kernel a usuario además de no necesitar una llamada al sistema por cada paquete como ocurre con `recvfrom()` por lo que el rendimiento de sistema de captura aumentaría.

Por último, el sensor utilizado en las pruebas aplica un procesamiento software a cada uno de los paquetes. Se puede elegir diferente grado de complejidad de análisis software para estudiar el impacto que tiene la carga de análisis en el rendimiento del sensor en la captura. Lógicamente, cuanto más tiempo dure el análisis menor será el tiempo disponible para la captura y las pérdidas que se producen aumentarán.

4.4 Análisis de los resultados

Considerando todos los tiempos indicados anteriormente, el tiempo medio que el SO emplea para capturar un paquete y ponerlo en una zona de memoria de usuario es la suma de las etapas B, C y D. La etapa A no entra en este cálculo porque es realizado por el motor DMA sin intervención de la CPU. La siguiente tabla resume los tiempos:

Tabla 5: Tiempos de cada etapa de la captura.

	A	B + C	D	Total
Ciclos	-	5.715	4.790	10,505
µs	-	2,37	2,01	4,38

Este tiempo es un tiempo mínimo ya que no se han tenido en cuenta los cambios de contextos debidos al planificador, ni el tiempo de *hardirq*, ni el tiempo que pasa desde la *hardirq* hasta la planificación de la *softirq*. Sin embargo, este dato nos sirve para establecer dónde está el límite superior de captura; dado que por cada paquete se emplean como mínimo 4,38 microsegundos, el número máximo de paquetes que un sistema de captura de las características del expuesto puede analizar sin pérdidas queda establecido en 228.310 paquetes por segundo.

El tiempo de *hardirq* se tuvo en cuenta a la hora de hacer las medidas sobre el kernel y su valor se expuso en el apartado anterior. Sin embargo, este tiempo no puede sumarse directamente en el cálculo debido a que es un tiempo compartido entre todos los paquetes que se capturen durante la *softirq*. No se ha tenido en cuenta porque es muy bajo y se puede despreciar.

Teniendo en cuenta todos los resultados de las pruebas realizadas, se observa que en una máquina Linux con las características expuestas en el apartado 5.1 y realizando capturas con un proceso a nivel de usuario, se empiezan a detectar pérdidas de paquetes debidas a problemas de agotamiento de recursos computacionales cuando la velocidad de los paquetes se acerca a 228.310 paquetes por segundo.

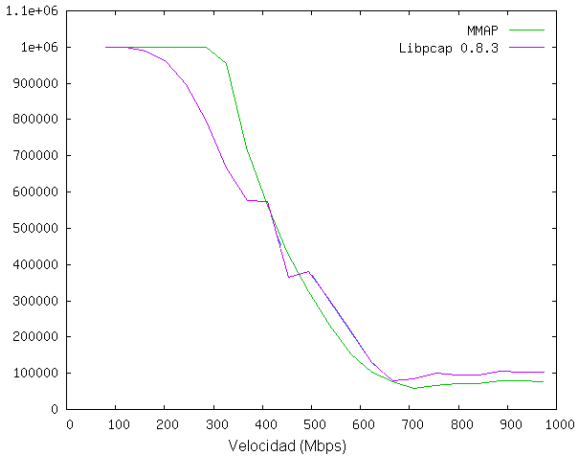


Figura 4: Mejoras con mmap (Y: paquetes capturado, X: velocidad inyección)

Realmente, en el laboratorio, las pérdidas empiezan mucho antes debido a las simplificaciones que se han realizado. Hay que destacar que éstos son los tiempos de captura suponiendo que la carga de análisis es nula, es decir, el caso mejor. Si el proceso de usuario decide manipular el paquete de alguna forma, por

ejemplo, para detectar intrusos decide buscar firmas de ataques en los paquetes sospechosos, el tratamiento de proceso (sección 3.2) consumirá más recursos del sistema reduciéndose el tiempo libre para el tratamiento de kernel. Al tener menos ciclos destinados a la captura las pérdidas aparecerán a velocidades más bajas.

Lógicamente, las etapas en las que se gasta más tiempo en la captura son la B, C y D. De todas ellas la D es evitable empleando mecanismos `mmap` [7] que mapean la memoria del kernel en espacio de usuario permitiendo el acceso directo de las aplicaciones a los paquetes almacenados en la memoria del kernel. Evitando esta última copia el rendimiento general del sistema aumenta (ver Figura 4). Las ordenadas representan el número de paquetes capturados mientras que las abcisas muestran la velocidad de inyección hasta 1.000 Mbps. En la figura se observa claramente como empleando `mmap` se capturan muchos más paquetes que con Libpcap especialmente antes de que el sistema se sature. Con el sistema saturado, las pérdidas utilizando ambas técnicas son similares.

Por otro lado, se ha identificado otra problemática menos conocida que los mecanismos `mmap` anteriormente comentados. El kernel durante una *softirq* es capaz de capturar varios paquetes. Por cada paquete se gastan una media de 5.700 ciclos repartidos en 2.900 para la copia desde la zona DMA y 1.550 para clonar el paquete. El punto de la captura en el que se pierden más paquetes es la cola del socket ya que, cuando el kernel intenta insertar un paquete en esta cola y ésta está llena, ese paquete será eliminado. Esta eliminación no sólo acarrea la pérdida del paquete sino que también es un consumo de recursos innecesario (concretamente unos 2.900 ciclos por paquete eliminado en esta cola). Por tanto, si estos ciclos malgastados fueran utilizados para la captura de paquetes que realmente vayan a ser analizados por el socket, el ratio de pérdidas disminuiría.

5 Conclusiones

Las directrices de diseño del mecanismo de captura de datos de un sistema operativo convencional son diferentes de las directrices de un sistema dedicado exclusivamente a la captura pasiva de tráfico en redes de alta velocidad. Por esta razón, la captura de datos no es óptima y puede ser mejorada. En este artículo se ha demostrado que, actualmente, la captura de datos basada en un sistema Linux convencional tiene carencias graves que no permiten alcanzar capacidades de análisis de tráfico adecuadas para segmentos de alta velocidad con tráfico comprometido.

Por otro lado, se ha demostrado también que, a velocidades de Gigabit, la capacidad de la CPU y el correcto diseño del software de captura es un factor clave para mejorar el rendimiento de un sistema de captura de tráfico en red. El resto de elementos hardware que intervienen tienen un ancho de banda que impone menos restricciones que la CPU. Debido a esto, en nuestro grupo de investigación estamos trabajando en la utilización de arquitecturas multiprocesador para aumentar la capacidad de procesamiento de los sensores con unos resultados preliminares bastante prometedores.

Por último, el detallado estudio del recorrido de un paquete en el kernel de Linux permite establecer con precisión que las copias de espacio de kernel a usuario y del área de DMA a kernel son muy costosas para el procesador. Un buen diseño software debería evitar en lo posible estas copias para mejorar el rendimiento. Además, la cola de socket sólo sirve para hacer la captura resistente a tráfico de ráfagas intenso, pero en general introduce un retardo innecesario cuando el tráfico es sostenido.

Referencias

[1] GNU/Linux. http://www.kernel.org.

[2] Tarjetas DAG. http://www.endance.com.

[3] Proyecto Scampi. http://www.ist-scampi.org.

[4] J. Mogul y K. Ramakrisnan, “Eliminating Receive Livelock in an Interrupt-Driven Kernel”, en Proceedings of Usenix Annual Technical Conference, 1996.

[5] J. Salim, “Beyond Softnet”, Proceedings of Usenix Annual Technical Conference, 2001.

[6] Libpcap, http://www.tcpdump.org.

[7] MMAP, Phil Wood, http://public.lanl.gov/cpw y en la documentación del kernel.

[8] Renato John Recio, “Server I/O Networks Past, Present, and Future”, en ACM SIGCOMM, 2003.

[9] A. Rubbini and J. Corbet, ”Linux Device Drivers”, 2nd Edition, O’Reilly, 2001.

[10] L. Rizzo, “Device Polling Support for FreeBSD”, BSDConEurope Conference, en 2001.

[11] “PCI-X Addendum to the PCI Local Bus Specification Revision 1.0”. 22 de Septiembre de 1999. http://www.pcisig.com/specifications/pcix_20/pci_x

Una mejora del framework SNMP de equilibrio de carga para controlar los computadores de la WLAN en zonas de cobertura reducida

David Sánchez, Elsa M. Macías, Álvaro Suárez
Grupo de Arquitectura y Concurrencia (GAC)
Departamento de Ingeniería Telemática
Universidad de Las Palmas de Gran Canaria
Campus Universitario de Tafira, 35017. Las Palmas de Gran Canaria, España
{dsanchez, emacias, asuarez}@dit.ulpgc.es

Abstract*. A network formed by desktop and portable computers is a useful environment for doing parallel computing. In this infrastructure we implement Master/Slave parallel distributed programs which exhibit strict data dependences among iterations and parallel calculations inside an iteration. Due to the dynamic behaviour of portable computers is necessary to keep in mind that they can move out of coverage area at any time. In a previous work, we developed a load balancing software framework based on Simple Network Management Protocol (SNMP) that considers the beacon strength in the portable computers for executing this kind of applications efficiently. However, when a portable computer is located in a limited coverage area, our framework considers that this resource is unavailable, and therefore it can't be used. For that reason, in this paper we present a mechanism that improves the performance of our framework SNMP, allowing us to use the portable computers while there is a wireless link.*

1 Introducción

Actualmente, el auge espectacular en la demanda de computadores portátiles con alto rendimiento, y la aparición de nuevos estándares de comunicación en la familia de protocolos IEEE 802.11, permiten una combinación efectiva de las redes de área local inalámbricas (WLAN) con las tradicionales redes de área local (LAN) para realizar computación paralela y distribuida [1], siendo ésta uno de los nuevos desafíos en los próximos años [2].

Nosotros hemos demostrado que la computación paralela en un entorno LAN-WLAN resulta eficiente para ejecutar aplicaciones Maestro/Esclavo con dependencias estrictas entre iteraciones, donde los cálculos paralelos son implementados en cada iteración [3]. Resulta evidente, que la heterogeneidad presente en este entorno de computación (diferentes arquitecturas, potencias de procesamiento y estándares de comunicación) lleva consigo una tarea ardua y difícil para ejecutar de forma eficiente de dichas aplicaciones. Si no se consideran técnicas de equilibrio de carga, algunos procesos del programa paralelo pueden caer en estados ociosos mientras otros están calculando.

En nuestro grupo de investigación hemos desarrollado una estrategia de equilibrio de carga [4][5] para aplicaciones Maestro/Esclavo que se desarrollan en un entorno de computación LAN-WLAN. Esta estrategia utiliza información acerca del rendimiento de los computadores y de la ejecución actual de cada proceso paralelo para estimar la distribución de datos adecuada. La recopilación de esta información tiene que ser obtenida en un segundo plano para no degradar el rendimiento de la aplicación paralela. En el trabajo realizado en [6], se utiliza el protocolo de gestión de red SNMP [7] para recoger de forma eficaz la información de rendimiento de los recursos fijos que forman parte del entorno de computación. En [5], desarrollamos una arquitectura software basada en el protocolo SNMP que obtiene de forma eficiente los parámetros de rendimiento de los computadores del entorno de computación LAN-WLAN. Estos parámetros son utilizados para llevar a cabo el equilibrio de carga en presencia de computación y comunicación heterogénea. En [8] desarrollamos una biblioteca que implementa la combinación de la estrategia de equilibrio de carga con la arquitectura SNMP, y por lo tanto, abstrae al usuario del conocimiento y de la implementación de nuestra arquitectura.

En este tipo de entornos, el comportamiento dinámico de los computadores portátiles implica que se tengan que aplicar técnicas de control que consideren los parámetros sobre la calidad del enlace inalámbrico y la energía de la batería. Dicha consideración viene determinada por el hecho de que no se debe esperar por resultados que no van a llegar, ya sea porque el recurso está fuera de cobertura, o porque la energía remanente en la batería no sea suficiente para finalizar la iteración en curso. En [9] nosotros modificamos nuestro *framework* SNMP para ejecutar eficientemente aplicaciones paralelas donde la estrategia de equilibrio de carga tiene en cuenta el nivel de potencia del enlace inalámbrico y la energía de la batería de los computadores portátiles. En dicho trabajo, cuando un recurso de computación está

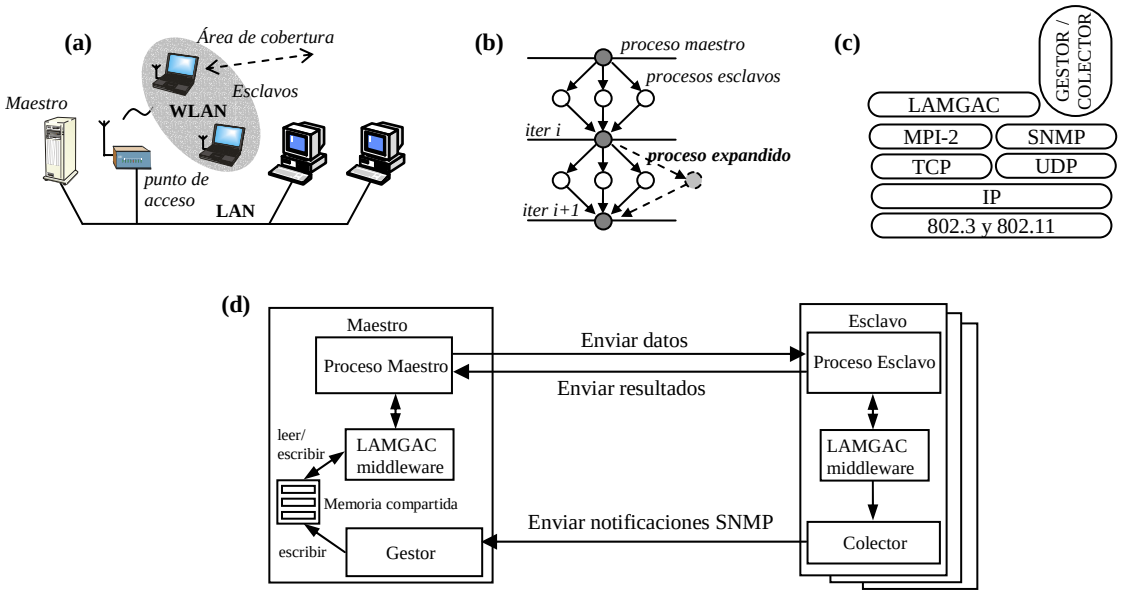


Fig. 1. a) Arquitectura hardware, b) Dependencias en aplicaciones Maestro/Esclavo, c) Arquitectura software, d) Interacción entre el framework SNMP y el programa paralelo

situado en una zona de cobertura reducida, la estrategia de equilibrio de carga informa, a la aplicación paralela, de que no se debe enviar ni recibir datos de dicho recurso porque puede desaparecer de forma inesperada. El recurso no participa en los cálculos mientras la calidad del enlace inalámbrico que recibe sea inferior a un cierto umbral. Como resultado, la potencia de cálculo del entorno se reduce, a pesar de que el proceso esclavo puede seguir realizando los cálculos y enviar los resultados obtenidos mientras hay conexión. El recurso permanece en una zona de cobertura reducida sin que por ello su intención sea abandonar el entorno de computación.

La nueva contribución que nosotros presentamos en este artículo es la extensión del *framework* SNMP para determinar de forma precisa los computadores de la WLAN que están accesibles, a nivel de comunicación y respecto al computador maestro, aún estando situados en una zona de cobertura reducida. De esta forma la potencia de cálculo del entorno se puede utilizar por completo mientras exista conexión con los computadores portátiles. Además, se propone un mecanismo de recepción controlada de datos que utilice la información suministrada por la arquitectura software para recibir datos sólo de aquellos computadores portátiles que estén accesibles.

El resto del artículo está organizado como sigue. En la sección 2, se describe de forma breve la arquitectura utilizada para ejecutar aplicaciones Maestro/Esclavo en un entorno LAN-WLAN, así como las principales funciones de la biblioteca diseñada. En la sección 3, se explica la implementación de la mejora introducida en la arquitectura, así como el modelo de recepción controlada que se debe utilizar. A continuación, en la sección 4, se muestran algunos resultados experimentales obtenidos. Por último, se presentan las conclusiones y el trabajo futuro.

2 Trabajo Previo

En la figura 1.a se muestra una combinación de la arquitectura LAN-WLAN. Por un lado, hay un computador maestro, que se encarga de distribuir la carga a los esclavos y es capaz de comunicarse con computadores situados en una LAN y en una WLAN (a través de un punto de acceso). Por otro lado, los computadores esclavos pueden pertenecer a una LAN o una WLAN. Los computadores de la LAN se consideran fijos y siempre están accesibles. Sin embargo, los computadores de la WLAN pueden cambiar su localización geográfica pudiendo entrar o salir del área de cobertura en tiempo de ejecución.

Nosotros consideramos aplicaciones paralelas en las cuales el proceso maestro distribuye, en cada iteración, una determinada cantidad de carga a cada proceso, figura 1.b. La distribución de datos la realiza el proceso maestro debido a que los procesos expandidos en tiempo de ejecución en los computadores portátiles sólo pueden comunicarse con éste. Cuando cada proceso finaliza sus cálculos, estos envían los resultados al proceso maestro. Este proceso tiene que recibir los resultados de todos los procesos esclavos antes de enviarles nuevos datos en la próxima iteración, debido a las dependencias de datos entre iteraciones en el tipo de aplicaciones consideradas.

En la figura 1.c se muestra la arquitectura software utilizada para desarrollar nuestro mecanismo de equilibrio de carga. Por un lado, utilizamos nuestro *middleware* LAMGAC [3], en su versión original,

basado en MPI-2 [10]. Este software nos permite gestionar la expansión dinámica de procesos esclavos en computadores portátiles que entran y salen de cobertura durante la ejecución de aplicaciones paralelas. La variación del número de procesos en los computadores portátiles puede ser controlada en cada iteración. Además, debido a las características heterogéneas del entorno y al tipo de aplicaciones que se ejecutan, es necesario aplicar algún mecanismo de equilibrio de carga para evitar estados ociosos en los procesadores más rápidos. Por este motivo, LAMGAC ha sido extendido en [5] y en [9] para implementar un mecanismo de equilibrio de carga efectivo en entornos de computación y comunicación heterogéneos. A continuación, en la tabla 1, se muestran las principales funciones de LAMGAC.

Tabla 1. Principales funciones de LAMGAC

Funciones	***Descripción***
LAMGAC_Update	Actualiza el número de procesos paralelos que se ejecutan en los computadores fijos y portátiles
LAMGAC_Balance	Estima la cantidad de datos que tiene que ser enviada a los procesos esclavos para alcanzar el equilibrio en los tiempos de ejecución
LAMGAC_ItestBattery_beacon	Averigua cuáles son los procesos ubicados en los computadores portátiles que están situados en una zona de cobertura reducida o cuya batería se agotará en breve
LAMGAC_Store_info	Almacena la cantidad de datos procesados por cada proceso y el tiempo empleado por éstos

Por otro lado, para llevar a cabo un equilibrio de carga eficiente se necesita información de rendimiento de los computadores. En este sentido, nosotros hemos diseñado un *framework* basado en SNMP para obtener dicha información de forma efectiva [5][9]. Debido a que la arquitectura diseñada debe consumir la cantidad mínima de recursos de cada computador, y no se necesita seguridad en las comunicaciones, se utiliza la versión SNMPv2C. A continuación, resumimos la interacción entre este *framework* y el programa paralelo, figura 1.d. En cada recurso se ejecuta un agente SNMP extendido, denominado *Colector*, que monitoriza algunos parámetros de rendimiento del computador, como son: la carga del procesador, latencia de comunicación, nivel de batería, calidad del enlace inalámbrico, etc. Cuando un evento significativo relacionado con estos parámetros ocurre (la calidad del enlace inalámbrico o nivel de batería está por debajo de un umbral, etc) el agente envía una notificación *(SNMP PDU-InformRequest)* a otro proceso SNMP, denominado *Gestor*, ubicado en el computador maestro. Este último proceso se encarga de obtener la información de rendimiento adjunta a la notificación recibida desde cada agente y, almacenarla en una zona de memoria compartida que es accedida por las funciones de la biblioteca LAMGAC. Esta información se utiliza para estimar la cantidad adecuada de carga a enviar a cada proceso cuando se invoca la función *LAMGAC_Balance()*, y para conocer los computadores con problemas de enlace inalámbrico y/o batería cuando se invoca la función *LAMGAC_ItestBattery_beacon()*.

3 Mecanismo de Mejora

En primer lugar, en esta sección se detalla la implementación previa a la mejora del *framework* SNMP en lo que respecta al control de los computadores portátiles. A continuación se presenta la mejora introducida para controlar de forma precisa los recursos que están en zonas de cobertura reducida, así como el esquema de recepción controlada.

3.1 Estado Actual

Como se explicó en la sección anterior, el proceso *Colector*, situado en los computadores esclavos, monitoriza diversos parámetros referentes al rendimiento de los computadores. En concreto, nuestro mecanismo de mejora se centra en el parámetro que refleja la calidad del enlace inalámbrico o nivel de potencia: el parámetro *lbLinkLevel* definido en la base de datos de información de gestión, LBGAC-MIB [9]. A continuación se detalla el funcionamiento del *framework* respecto a este parámetro.

El proceso *Colector* monitoriza periódicamente el parámetro *lbLinkLevel*, y envía dicho valor al proceso *Gestor* cuando se produce alguna de las siguientes situaciones:

- Comienza la ejecución de un nuevo proceso paralelo en el computador.
- Descenso continuo de la potencia de señal inalámbrica durante un intervalo de tiempo.
- La potencia de señal inalámbrica es inferior a un umbral establecido.
- Hay nivel de señal después de haber estado un periodo de tiempo fuera de cobertura.

Cada una de estas situaciones se corresponde con el envío de una notificación asíncrona, las cuales están definidas en LBGAC-MIB, y se indican en la tabla 2.

Por parte del *Gestor*, una vez que se extrae de la notificación el parámetro *lbLinkLevel*, éste se compara con un umbral fijado por el usuario. Si dicho umbral es superior al nivel de potencia indicado, el

computador se considera situado en un área de cobertura limitado, y por tanto, se declara el recurso como no accesible. Una llamada a la función *LAMGAC_ItestBattery_Beacon()* indicaría el rango del proceso que se ejecuta en dicho computador. El proceso maestro no debe considerar ese proceso en las sucesivas distribuciones de datos mientras la calidad del enlace inalámbrico que recibe el computador portátil no esté por encima del umbral.

Tabla 2. Notificaciones que incluyen el objeto lbLinkLevel

Notificaciones	*Descripción*
lbnAppStart	Se envía cuando un nuevo proceso paralelo comienza su ejecución en el recurso
lbnDescLink	Se envía cuando se detecta un descenso consecutivo en la potencia de la señal
lbnCriticalArea	Se envía cuando el nivel de señal está por debajo de un cierto umbral
lbnUpLink	Se envía cuando hay nivel de señal después de haber estado fuera de cobertura

El valor del umbral presenta un claro compromiso. Si se toma un valor alto, se reduce el área de movilidad de los recursos ya que éstos se consideran inalcanzables cuando la potencia de la señal disminuye levemente y es inferior al umbral (aunque pueda existir conexión), con el resultado de no aprovechar la potencia de cálculo del computador. Sin embargo, si la elección del umbral es un valor muy bajo puede darse el caso que el recurso quede fuera de cobertura de forma inesperada sin dar a lugar a enviar una notificación del tipo *lbnCriticalArea*, con la consecuencia pérdida de control sobre el recurso. Esta última situación puede suponer esperas indefinidas en la recepción de datos, con el consecuente bloqueo del proceso maestro esperando por datos que no van a llegar desde el proceso esclavo.

En este sentido, la mejora que se explica a continuación va en la línea de acotar el rango de incertidumbre que existe al fijar el umbral de cobertura, ya que el recurso sólo se declara inaccesible cuando se determina de forma precisa que no existe comunicación entre éste y el computador maestro.

3.2 Control de los Computadores Portátiles

El nuevo mecanismo que se presenta en esta sección consiste en la ampliación del *framework* SNMP. En la figura 2 se muestra un esquema de funcionamiento. Esta ampliación se basa en la creación, por parte del proceso *Gestor*, de una hebra de ejecución *(thread)* cuando el umbral de cobertura es superior al nivel de potencia de la señal inalámbrica que recibe un computador portátil (hebra *SnmpPing* en la figura 2). El proceso *Gestor* crea tantas hebras como computadores estén situados en una zona de cobertura limitada. El proceso *Gestor* conoce la calidad del enlace inalámbrico que recibe cada computador portátil a través de las notificaciones (tabla 2) enviadas por cada proceso *Colector*.

La función de las hebras de ejecución es monitorizar el estado del enlace de comunicación entre el computador maestro y los computadores portátiles mientras estos estén ubicados en un área de cobertura reducida. Para ello, cada hebra, mediante una operación *Get-Request* del protocolo SNMP, consulta la potencia de señal recibida en el computador que envió alguna de las notificaciones descritas en la tabla 2, es decir, se consulta al agente *Colector* sobre el objeto *lbLinkLevel*. Cada hebra permanece en ejecución mientras el computador portátil esté ubicado en el área de cobertura limitada. Por lo tanto, la hebra finaliza cuando el nivel de potencia de la señal inalámbrica está por encima del umbral (cobertura aceptable) o cuando no hay enlace de comunicación entre el computador maestro y el esclavo (expira el *timeout* de varias operaciones *Get-Request* consecutivas). Una vez que se den las condiciones que provocan la finalización de la hebra de ejecución, se escribe en la memoria compartida por la biblioteca LAMGAC y el proceso *Gestor* el estado del computador portátil: accesible o no. Esta información es consultada por la función de control de batería y enlace (*LAMGAC_ItestBattery_beacon()*) para informar al proceso maestro qué procesos se están ejecutando en los computadores no disponibles.

La implementación de este mecanismo de mejora modifica la definición de la información retornada por la función *LAMGAC_ItestBattery_beacon()* de la siguiente forma (en cursiva resaltamos las modificación de la definición con respecto a la mostrada en tabla 1): "Averigua cuáles son los procesos ubicados en los computadores portátiles que *no tienen conexión de red inalámbrica con el computador maestro* o cuya batería se agotará en breve".

Por otro lado, debido al comportamiento dinámico del canal de comunicaciones inalámbrico, la carga del procesador y la localización del computador portátil, el tiempo transcurrido desde que la consulta SNMP se realiza hasta que la respuesta se recibe puede variar de forma considerable de un instante a otro. Este tiempo representa un serio compromiso en la elección del valor del *timeout* de la operación de consulta. Un valor elevado provoca un tiempo de espera alto cuando existe un fallo en la comunicación (fuera de cobertura, fallo en el canal, etc) y, por lo tanto, la función de control de batería y enlace pueden indicar que existe una conexión física cuando realmente no la hay (se invocan las funciones cuando aún no ha expirado el *timeout* de la operación de

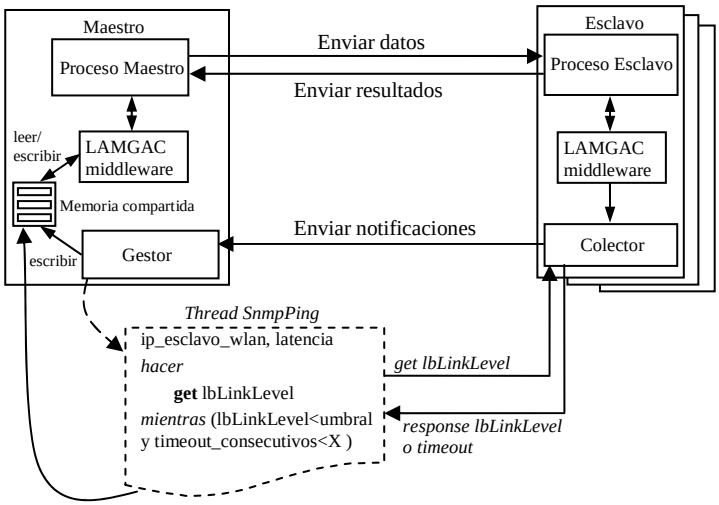


Fig 2. Framework SNMP modificado con hebra de ejecución SnmpPing

consulta y la información almacenada en la memoria compartida no es consistente con la situación actual). Por otro lado, si se opta por elegir un valor demasiado bajo, cualquier situación de congestión en el canal puede provocar que dicha función indique que no hay conexión cuando realmente si la hay (el *timeout* de la operación de consulta expira sin dar a lugar a la llegada de la respuesta enviada por el proceso *Colector*). Teniendo en cuenta este compromiso, el valor apropiado para el *timeout* debe ser ligeramente superior al tiempo consumido al realizar una consulta con éxito. Matemáticamente, el tiempo máximo permitido para llevar a cabo una operación de consulta al computador *i* (c_i) se calcula como:

$$tout(c_i) = 2 \times t_{lat}(c_i) + \frac{enviados + recibidos}{B} + t_{get}(c_i)$$

donde:

- $t_{lat}(c_i)$ es la latencia de comunicación entre el computador maestro y el computador portátil. Por simplicidad, asumimos que este valor es siempre igual en ambos sentidos de comunicación, y que su valor permanece constante. Este valor y el parámetro *B* son calculados por el agente *Colector* cuando el demonio *lamd* (distribución LAM-MPI) se inicia, y son enviados al proceso *Gestor* mediante la notificación *lbnAppStart.*
- *enviados* es el tamaño en bytes del paquete generado en la operación de consulta. En nuestro entorno de computación su valor es de 92 bytes.
- *recibidos* es el tamaño en bytes del paquete generado en la operación de respuesta a la consulta. En nuestro entorno de computación su valor es de 93 bytes.
- *B* es la velocidad de transmisión entre el computador maestro y el computador esclavo.
- $t_{get}(c_i)$ es el tiempo consumido por el agente *Colector* para decodificar la consulta, ejecutarla y devolver el resultado. Este valor depende de varios factores, como son: velocidad del procesador, tamaño de memoria, carga del sistema, etc. Por lo tanto, para calcular el valor de $tout(c_i)$, este valor se ha estimado de forma empírica para el entorno y condiciones en las que trabajamos (especificado en la tabla 3 de la sección 4). Se han realizado muchas medidas sobre el tiempo transcurrido al realizar una consulta *Get-Request* con éxito, y el valor medio obtenido es aproximadamente 0.7 ms.

Por lo tanto, los datos que tiene que comunicar el proceso *Gestor* a la hebra de ejecución para que ésta pueda realizar la consulta SNMP con el valor de *timeout* apropiado, son: la dirección IP del computador portátil y la latencia de comunicación entre este último y el computador maestro.

Por otro lado, sabiendo que SNMP utiliza el protocolo de transporte no fiable UDP y que la cantidad de bytes enviados y recibidos es pequeña, podemos asumir que el tráfico generado debido a la consulta afecta de forma mínima al rendimiento de la red [11].

3.3 Esquema de Recepción Controlada

Una vez iniciada la recepción de datos por parte del proceso maestro puede que algún computador portátil se sitúe fuera del área de cobertura, y por lo tanto, no

pueda enviar sus datos. Si esto ocurre, el proceso maestro espera por los datos calculados en dicho computador el tiempo que éste tarde en incorporarse al área de cobertura o indefinidamente si no regresa a la WLAN. Esta espera es perjudicial para la aplicación paralela, ya que introduce desequilibrios en los tiempos de ejecución de los procesos. Para solucionarlo, se propone un esquema de recepción controlada para la recepción de datos, basado en funciones de LAMGAC. A continuación, se explica en detalle dicho esquema, el cual se presenta en la figura 3.

```
LAMGAC_ItestBattery_beacon (…);
∀ procesos disponibles
  // Se inicia la recepción de datos no bloqueante

mientras queden operaciones de recepción por completar
  // Comprobar operaciones completadas
  ∀ nuevas operaciones de recepciones completadas
      LAMGAC_Store_info (…);

  LAMGAC_ItestBattery_beacon (…);
  ∀ nuevos procesos no disponibles
      // Cancelar la recepción de sus datos
fin mientras
```

Fig 3. Esquema de recepción controlada

Antes de iniciar la recepción de datos, se realiza una llamada a la función *LAMGAC_ItestBattery_beacon()* para conocer si hay algún computador que actualmente están inalcanzable. Si existen algún recurso con el que no es posible la comunicación desde el computador maestro, entonces no se implementa la operación de recepción para el proceso que se ejecuta en ese computador. Una vez iniciadas las recepciones no bloqueantes de los procesos esclavos que se desarrollan en los computadores accesibles, el proceso maestro se queda esperando por los resultados. Durante dicha espera, continuamente se comprueba si ha completado alguna de las operaciones de recepción iniciadas y si hay algún nuevo recurso con problemas de cobertura. En el caso de que alguna operación se complete, se llama a la función *LAMGAC_Store_info()* para almacenar en la memoria compartida los datos referentes al rendimiento del proceso cuya operación de recepción se acaba de completar. Estos datos son utilizados por el mecanismo de equilibrio de carga. En el caso de que existan nuevos procesos con problemas de cobertura, se cancela la recepción de éstos para no provocar esperas indefinidas por resultados que nunca van a llegar. El programador de la aplicación tiene que considerar este hecho para calcular los resultados que no pudieron ser devueltos por los procesos cancelados. Este bucle se repite continuamente hasta que no existan datos por recibir.

4 Resultados Experimentales

La mejora realizada en la arquitectura SNMP aporta claras ventajas en la ejecución de la aplicación paralela, ya que los procesos paralelos que se desarrollan en los computadores portátiles son considerados por la aplicación mientras exista comunicación con el computador maestro. Sin embargo, la ejecución de la hebra SnmpPing implica una sobrecarga en el entorno de computación. En esta sección se presenta cómo afecta la mejora realizada en el *framework* SNMP en el tiempo de ejecución global de la aplicación paralela.

Para ello, se han realizado varios experimentos con la versión previa de la arquitectura SNMP y la nueva mejora aportada en este artículo. El entorno de computación utilizado está formado por una red de computadores que combina segmentos de red IEEE 802.3 e IEEE 802.11. Las características de los computadores se indican en la tabla 3, los cuales utilizan la librería de paso de mensajes MPI-2 bajo el sistema operativo Linux. Cada simulación se repitió 10 veces obteniendo una desviación típica reducida. La aplicación secuencial se ejecutó en el procesador más rápido.

Tabla 3. Características de los recursos

Procesador / Memoria	***Red (Mbps)***
PIV 2.4Ghz/512 MB (maestro)	100
PIV 2.4Ghz/512 MB	100
PIV 2.4Ghz/512 MB (portátil)	11
PIV 2.4Ghz/512 MB (portátil)	11
PIII 450/128MB	100

Para llevar a cabo los experimentos, se ha utilizado como aplicación paralela una herramienta de Codiseño Hw/Sw, explicada con detalle en [5], utilizando el modelo de programación para equilibrio de carga especificado en [8]. Esta herramienta calcula la mejor combinación de recursos Hw/Sw, que cumple una serie de restricciones, para implementar un sistema de reconocimiento de voz. La especificación del sistema viene dada en el lenguaje de descripción VHDL. Esta herramienta posee dos aplicaciones: una fase de estimación y otra de particionado. Para realizar los experimentos se ha utilizado sólo la fase de estimación.

Antes de cada distribución de datos, el proceso maestro calcula, con un procedimiento recursivo, todas las combinaciones posibles de recursos Hw/Sw para implementar un proceso VHDL del sistema. Como este procedimiento no está paralelizado, el *speedup* de la aplicación se reduce considerablemente (figura 4.b). Después de calcular las combinaciones, el proceso maestro distribuye a los procesos esclavos un rango de combinaciones de recursos para estimar el costo de la implementación de cada una.

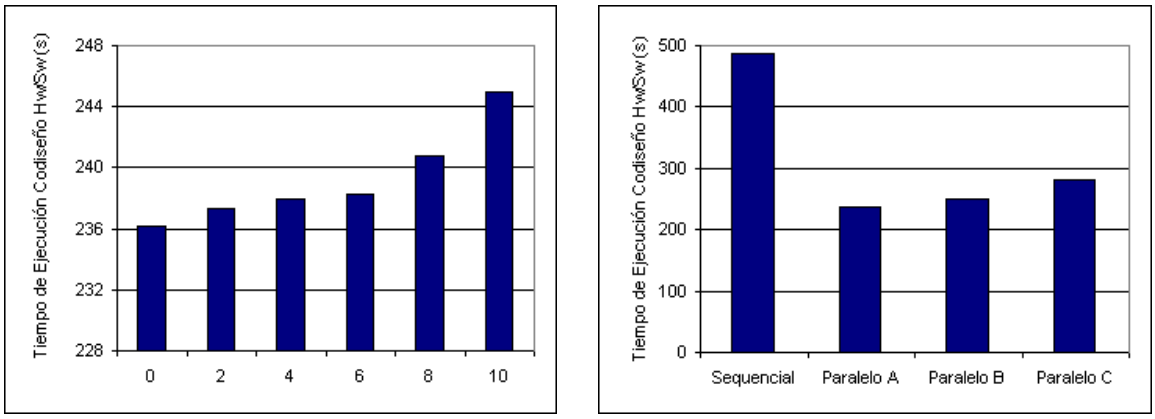


Fig 4. a) Sobrecarga de la hebra SnmpPing, b) Mejora del tiempo de ejecución

4.1 Sobrecarga de la Hebra

La ejecución de la hebra SnmpPing introduce sobrecarga en el computador maestro, en el computador portátil y en la red. Por lo tanto, es necesario estudiar la sobrecarga que introduce en el tiempo de ejecución de la aplicación paralela. La sobrecarga se debe a:

- Computador maestro: crear la hebra, construir la operación *Get-Request*, enviar la consulta al proceso *Colector* y, recibir y almacenar los datos en la memoria compartida.
- Computador portátil: decodificar la consulta SNMP, obtener la información y enviarla al proceso *Gestor*.
- Red: tráfico de paquetes UDP para las tramas de consulta y respuesta.

El primer y segundo punto depende del rendimiento y la carga actual de los computadores implicados. El tercero depende del rendimiento de la red, como puede ser: la latencia de comunicación y la tasa de transferencia. Además, también se ve afectado por las características especiales de las redes inalámbricas (alta congestión, cobertura, medio compartido, etc). Teniendo en cuenta estas consideraciones, los resultados presentados en esta sección pueden variar de forma considerable de un entorno de computación a otro.

Para evaluar la sobrecarga que introduce la hebra SnmpPing, se han realizado varios experimentos situándonos en el peor escenario posible, es decir, se ha forzado la ejecución de una hebra desde el comienzo de la aplicación paralela por cada uno de los recursos portátiles que intervienen en la aplicación. Debido al número limitado de computadores portátiles disponibles, igual a dos, la creación de un número de hebras superior a dos se realiza replicando las hebras creadas.

La figura 4.a muestra el tiempo de ejecución medio de la fase de estimación de la herramienta de Codiseño Hw/Sw en función del número de hebras ejecutadas. Como se puede apreciar, el tiempo de ejecución de la aplicación paralela sigue aproximadamente una relación lineal con el número de hebras, aumentando sobre nueve segundos cuando se ejecutan diez hebras durante toda la ejecución de la aplicación paralela. En cualquier caso, este incremento tiene una influencia mínima en el tiempo de ejecución global y, por lo tanto, la sobrecarga es despreciable. Además, en una situación real, las hebras se crean y destruyen dinámicamente si se producen las condiciones comentadas en el apartado 3.2 (sólo cuando el computador está en una zona de cobertura reducida), por lo que en la práctica es más que probable que la sobrecarga que introducen estas hebras sea inferior a la que nosotros forzamos en nuestra simulación.

4.2 Mejora del Tiempo de Ejecución

Para demostrar la mejora del tiempo de ejecución cuando se utiliza la aportación presentada en este artículo, se han realizado varias simulaciones donde los computadores portátiles están localizados dentro del área de cobertura durante algunas iteraciones. En concreto, la aplicación paralela tiene ocho iteraciones, y durante las cuatro primeras los computadores portátiles se ubicaron en el área de cobertura limitada.

En la figura 4.b se muestran los resultados. El experimento etiquetado como *Paralelo A* indica el tiempo medio de ejecución de la aplicación paralela cuando se utiliza la mejora aportada en este artículo. Los experimentos etiquetados como *Paralelo B* y *Paralelo C* representan el tiempo medio de ejecución cuando no se utiliza la nueva contribución y, además, hay uno y dos computadores en el área de cobertura limitada, respectivamente. Como se puede apreciar, el tiempo de ejecución global se reduce cuando se utiliza la contribución de este artículo. Con respecto al experimento *Paralelo C*, el tiempo se reduce aproximadamente en 50 segundos, lo que representa una disminución del 18%.

5 Conclusiones y Trabajo Futuro

En este artículo se ha presentado una mejora sobre una arquitectura SNMP desarrollada en nuestro grupo de investigación para equilibrar la carga en aplicaciones paralelas que se desarrollan en entornos heterogéneos LAN-WLAN. La mejora consiste en aprovechar la capacidad de procesamiento de los computadores portátiles mientras exista un enlace de comunicación entre el proceso maestro y éstos. Esta mejora se ve complementada con la propuesta de una recepción de datos controlada. Además, hay que añadir que esta mejora no introduce ningún cambio en la utilización de las funciones de LAMGAC definidas en trabajos previos, por lo que es transparente al usuario.

Respecto al trabajo futuro, queremos diseñar un esquema para detectar aquel computador portátil que ha desaparecido de forma inesperada del área de cobertura sin darle tiempo a notificar este hecho al proceso maestro. Además, en éste y en trabajos previos consideramos constante la latencia de red. Debido a las características de las redes WLAN, este parámetro puede variar fuertemente de un instante a otro. Por lo tanto, queremos modificar la arquitectura SNMP para estimar el valor de la latencia de forma periódica, de tal manera que afecte de forma mínima al tráfico generado por otras aplicaciones.

Agradecimientos

Este trabajo está subvencionado por la Consejería de Educación, Cultura y Deportes del Gobierno de Canarias (PI:164/2004).

Referencias

[1] Cheng, L. Wanchoo, A., Marsic, I., "Hybrid Cluster Computing with Mobile Objects", 4th IEEE International Conference on High Performance Computing in Asia-Pacific. Beijin, China (2000) 909-914.

[2] A. Zomaya. "Mobile Computing: Opportunities for Parallel Algorithms Research", Proceedings of the International Parallel and Distributed Processing Symposium, USA, 2002, 144-147.

[3] E. Macías, A. Suárez. "Solving Engineering Applications with LAMGAC over MPI-2", Proceedings of 9th European PVM/MPI. LNCS 2474, Springer Verlag. Linz, Austria, 2002, 130-137.

[4] D. Sánchez, E. Macías, A. Suárez. "Effective Load Balancing on a LAN-WLAN Cluster", Proceedings of Parallel and Distributed Processing, Techniques and Applications, Vol. I, CSREA, Las Vegas, USA, 2003, pp. 473-479.

[5] D. Sánchez, E. Macías, A. Suárez. "Anticipating Performance Information of Newly Portable Computers on the WLAN for Load Balancing". Proceedings of 5th Parallel Processing and Applied Mathematics, LNCS 3019. Springer-Verlag, Czestochowa, Poland, 2003, 946-953.

[6] R. Busby, M. Nielsen, D. Andresen, "Enhancing NWS for use in an SNMP Managed Internetwork", Proceedings of International Parallel and Distributed Processing Symposium, Cancún, Mexico, 2000, 506-511.

[7] M. Subramanian, *Network management: principles and practice*, Addison-Wesley, USA, 2000.

[8] D. Sánchez, E. Macías, A. Suárez. "A library for Load Balancing in master/Slave Applicattions on a LAN-WLAN Environment". Proceedings of the 12th Euromicro Conference on Parallel, Distributed and Network-based Processing. A Coruña, España, 2004, 168-175.

[9] D. Sánchez, E. Macías, A. Suárez. "Load balancing Detecting Battery Energy Level and Wirelles Beacon Strength". Proceedings of the 15th IASTED International Conference on Parallel and Distributed Computing and Systems. Marina del Rey, USA, 2003, 268-273.

[10] William Gropp, Ewing Lusk, Rajeev Thakur, *Using MPI-2: advanced features of the message-passing interface*, Cambridge, Mass. MIT Press, 1999.

[11] M.G. Arranz, R. Agüero, L. Muñoz, P. Mähönen. "Behaviour of UDP-Based Applicationes over IEEE 802.11 Wireless Networks". Proceedings of 12th IEEE International Symposium on Personal Indoor and Mobile Radio Communication. San Diego, USA, 2001, vol II 72-77.

QoSM3. Herramienta de modelado de tráfico y tomografía de red para servicios de telemedicina

I. Martínez, A. Valero, E. Viruete, J. Fernández, J. García

Grupo de Tecnología de las Comunicaciones (GTC). Instituto de Investigación de Ingeniería en Aragón (I3A)
Centro Politécnico Superior (CPS). Universidad de Zaragoza (UZ).
Edificio Ada Byron. Campus Río Ebro. c/María de Luna 3, 50.018 – Zaragoza (Spain)
Teléfono: 976 76 19 45 Fax: 976 76 21 11 E-mail: imr@unizar.es

Abstract*. The wide development of multimedia clinical applications and the use of inter and intra-hospital communication networks require a specific analysis to increase healthcare services efficiency. In this paper we propose a processing toolbox (QoSM3) for technical evaluation of Quality of Service (QoS) traffic requirements in new healthcare services based on telemedicine. This tool consists of the multimedia service definition and the measurement and modelling processes which permit to analyse QoS requirements and to optimize application design regarding available network resources. The proposed methodology has been tested to evaluate real-time and store&forward medical services.*

1 Introducción

Las nuevas tecnologías han permitido que los servicios de telemedicina hayan experimentado un importante avance y desarrollo en los últimos años. Para extraer el máximo beneficio de estos nuevos servicios, resulta imprescindible definir una metodología precisa para caracterizar los requisitos planteados en la transmisión de la información y en la gestión de los recursos de red disponibles [1]. Además, es indispensable llevar a cabo su correcta evaluación incluyendo aspectos de eficiencia, aceptabilidad y usabilidad para que puedan incorporarse a los sistemas de salud en los diferentes escenarios asistenciales (entornos rurales, teleasistencia, asistencia domiciliaria, etc.) [2]-[3].

En este instante surge la necesidad de optimizar la calidad de servicio (*Quality of Service*, QoS) que se obtiene de dichos servicios de telemedicina [4]-[5]. Para ello, es crucial el estudio de dos aspectos: la naturaleza de la información biomédica a transmitir, y el comportamiento de las redes que la transportan. La información asociada a las aplicaciones médicas requiere un conocimiento detallado y una caracterización de modelos que las definan [6]. Igual que la variabilidad de las prestaciones (p.ej. las infraestructuras móviles) y la heterogeneidad de las interconexiones (p.ej. Internet) requieren medir y modelar las redes de interconexión [7]-[8].

En esta línea, una idea extendida consiste en que es posible gestionar y adecuar de forma adaptativa la transmisión de la información generada por las aplicaciones (*codecs*, tasa de transmisión y compresión, etc.) a los recursos de las redes que atraviesan. Esto permitiría mejorar la QoS de las comunicaciones e-sanitarias, buscando que sea óptima en cada momento [9]-[10].

Establecer modelos a partir de medidas de red significa monitorizar la red inferencialmente, con el objetivo de obtener información para la posterior actuación. Esta disciplina telemática se conoce como tomografía de red [11]. Existen diversas aplicaciones desarrolladas en esta monitorización inferencial y suelen obtener una caracterización de la topología de la red a partir de la estimación de valores de algoritmos de encaminamiento [12]. Otros trabajos se centran en obtener parámetros de QoS para modelar no solo la topología, sino los propios servicios generadores del tráfico [13].

Así, cuando se hace un estudio de modelado de tráfico para ofrecer QoS, es imprescindible contar con modelos de las fuentes y de las redes que van a componerlo: entender la dinámica del tráfico, y usar ese conocimiento en el diseño; el papel de ingeniero de teletráfico [14] sería de realimentación donde las medidas informan sobre su comportamiento y los criterios de QoS definen su funcionamiento. Este planteamiento estructuralista (a diferencia de los trabajos clásicos basados en modelos conductistas [15], que imitan las propiedades estadísticas del tráfico resultante sin tener en cuenta la generación del mismo) es más exacto [16] ya que, si el objetivo del modelado es entender la dinámica del mismo, ese conocimiento es el que debe emplearse para diseñar, gestionar, y controlar los servicios y redes existentes y futuros. Este proceso, además, intenta apoyarse en la parsimonia; es decir, que el modelo completo pueda ser definido mediante un conjunto reducido de parámetros (distribución probabilística, valores medios, desviaciones típicas, correlación, autosimilitud, etc.) que, por lo general, suelen denotar un significado físico (tiempo entre sucesos, duración de los mismos, etc.).

En este artículo, se propone una herramienta para el modelado del tráfico y la tomografía de red a partir de simulaciones y medidas experimentales, siguiendo requisitos de QoS en los nuevos servicios sanitarios. Estos modelos son específicos para telemedicina (ya que integran usuarios dispersos, con aplicaciones sanitarias de diversa naturaleza y redes hospitalarias heterogéneas), pero podrían ser válidos para cualquier escenario multimedia en que se exija QoS.

En la sección 2 se describe la topología básica del escenario de estudio, la metodología de evaluación seguida, y las variables definidas en la analítica. En la sección 3 se presenta la herramienta QoSM3 de caracterización y modelado. Esta herramienta tiene como objetivo la optimización del diseño de las aplicaciones e incluye diversas técnicas y algoritmos para calcular los diferentes parámetros de QoS. Los resultados obtenidos y su aplicación a la evaluación de los nuevos servicios de telemedicina sobre diferentes entornos de red se discuten en la sección 4.

2 Descripción de la herramienta

La implementación de nuevos servicios sanitarios basados en telemedicina requiere una evaluación técnica para estudiar su implementación bajo diferentes condiciones de red. Por ello, se ha desarrollado un proceso automatizado para la medida de los parámetros de QoS y el modelado del servicio multimedia basado en la herramienta *Service M3* (*Multimedia Measurement & Modelling*) [17]. *Service M3* incluye tres módulos: uno de definición multimedia que traduce los requerimientos clínicos en parámetros telemáticos; otro de medida que captura tanto el tráfico experimental -*RealM3*- como el simulado -*SimulatedM3*- en formato homogéneo; y un último módulo común de modelado (ver Fig. 1).

Este último módulo -*QoSM3*- recibe como entrada un fichero de trazas, obtenido a partir de las medidas o simulaciones previas, y genera un modelo completo del tráfico y de la red. Incluye una herramienta que permite representar gráficas y estadísticas en un formato portable que facilita la evaluación. De esta forma, la interpretación y comparación de resultados (no sólo entre diferentes pruebas de un mismo escenario, sino de la misma prueba en diferentes escenarios) permite caracterizar el comportamiento del servicio para evaluar el sistema completo. Todas estas utilidades conforman el paquete (*toolbox*) programado en C, que se divide en tres bloques:

2.1 QoSM3 Básico

Este bloque implementa un análisis básico, previo al diseño del servicio y con vistas a caracterizarlo, que describe sus aspectos elementales. Incluye el cálculo de las tasas medias y de pico, *SDR* y *PDR* en (1), para un intervalo de muestras *n* elegido, y la estimación de latencia, tiempo de transmisión, y retardos significativos incluidos en el Apéndice I (ver Fig. 2): retardo de transmisión ($\tau^{l}_{c.i}$), procesado ($p^{l}_{c.i}$), espera en cola ($q^{l}_{c.i}$), acceso al medio ($a^{l}_{c.i}$), propagación (d^{l}_{prop}), etc.

2.2 QoSM3 Aplicación

Este bloque contiene métodos de análisis y ajuste probabilístico para y modelar el comportamiento del tráfico observado [18]. Permite una validación estadística para una o varias conexiones simultáneas, pudiendo realizar medidas individuales o conjuntas (*cross-traffic*). Estudia el rafagueo de las fuentes sobre el escenario propuesto (módulo de ráfagas), y analiza el nivel de ocupación de los *buffers* intermedios, así como otros parámetros de red que se ven afectados por el comportamiento del tráfico que la atraviesa (módulo de flujo). Así, la conjunción de ambos módulos persigue una caracterización completa del tráfico observado que permitiera adecuar la QoS del servicio según los requisitos establecidos para telemedicina. Se divide en:

A. Módulo de ráfagas.
A partir de los parámetros de entrada, estima el tamaño máximo de ráfaga *MBS* en (2), y la tolerancia a ráfagas *BT* en (3). Además, para la caracterización, evolución y comportamiento de las fuentes de tráfico (aisladas y/o multiplexadas) calcula variables de utilidad mediante un proceso de dos fases:

- Caracterización de primer orden [19], calculando valores estadísticos como media μ y varianza σ^2, histogramas *h(i)* en (4), funciones de densidad *PDF* en (5), de distribución *CDF* en (6), etc.
- Caracterización de orden superior [20], implementando herramientas de análisis y cuantificación del rango de dependencia temporal de las muestras capturadas como la autocorrelación en (7) o el grado de subexponencialidad en (8).

Peak & Sustained Data Rate (PDR & SDR)	$PDR_i = \frac{s_i}{\Delta t_i} \rightarrow SDR_i^n = \frac{\sum_n s_i^n}{t_{i+n} - t_i}$	(1)
Maximum Burst Size (MBS)	$MBS = \left\lfloor 1 + \frac{BT}{T_s - T} \right\rfloor \quad with \quad \overline{PDR} = 1/T \;\; \overline{SCR} = 1/T_s$	(2)
Burst Tolerance (BT)	$BT = (MBS - 1) \cdot \left(\frac{1}{SDR} - \frac{1}{PDR}\right)$	(3)
Histogram (for an interval $i \in [r_{i-1}, r_i)$ between L_h)	$h(i) = \sum_{n=1}^{N} I_{[r_{i-1}, r_i)}(\Delta t_i[n]) \;\; with \;\; i = 1, \ldots, L_h$ $and \;\; I_{[r_{i-1}, r_i)}(\Delta t_i[n]) = \begin{cases} 1 & if \;\; r_{i-1} \le \Delta t_i[n] < r_i \\ 0 & other\ cases \end{cases}$	(4)
Probability Density Function (PDF)	$\hat{f}_{\Delta t}(\Delta t) = \sum_{i=1}^{L_h} I_{[r_{i-1}, r_i)}(\Delta t_i) \cdot \frac{h(i)}{N \cdot (r_i - r_{i-1})}$	(5)
Cumulative Distribution Function (CDF)	$\hat{F}_{\Delta t}(\Delta t) = \begin{cases} \sum_{i=1}^{L_h} I_{[r_{i-1}, r_i)}(\Delta t_i) \cdot \left(\frac{\sum_{j=1}^{i} h(j)}{N}\right) & si \;\; \Delta t_i < máx(\Delta t_i[n]) \\ 1 & si \;\; \Delta t_i < máx(\Delta t_i[n]) \end{cases}$	(6)
Autocorrelation	$\hat{R}_{\Delta t}(k) = \frac{1}{\sigma^2_{\Delta t}(N-k)} \sum_{i=1}^{N-k} (\Delta t_i - \mu)(\Delta t_{i+k} - \mu)$	(7)
Subexponenciality (*Hill estimator*)	$\hat{\alpha}(k) \propto \left[\frac{1}{k} \sum_{i=1}^{k-1} \mathrm{Log}\left(\frac{\Delta t[N-1]}{\Delta t[N-k]} \right) \right]^{-1}$	(8)
Utilization factor (normalized)	$\rho^* = \frac{\rho}{\rho_{máx}} = \frac{C_e / C}{C_{e_{máx}} / C} = \frac{C_e}{C_{e_{máx}}}$	(9)
BW Estimation (BE) $\left(with \;\; \alpha_k = \frac{2\tau - \Delta t_k}{2\tau + \Delta t_k}\right)$	$BE_i = \alpha_k \cdot BE_{i-1} + \frac{1}{2}(1 - \alpha_k)(BW_i + BW_{i-1})$	(10)

Tabla 1. Algunas expresiones teóricas usadas en QoSM3.

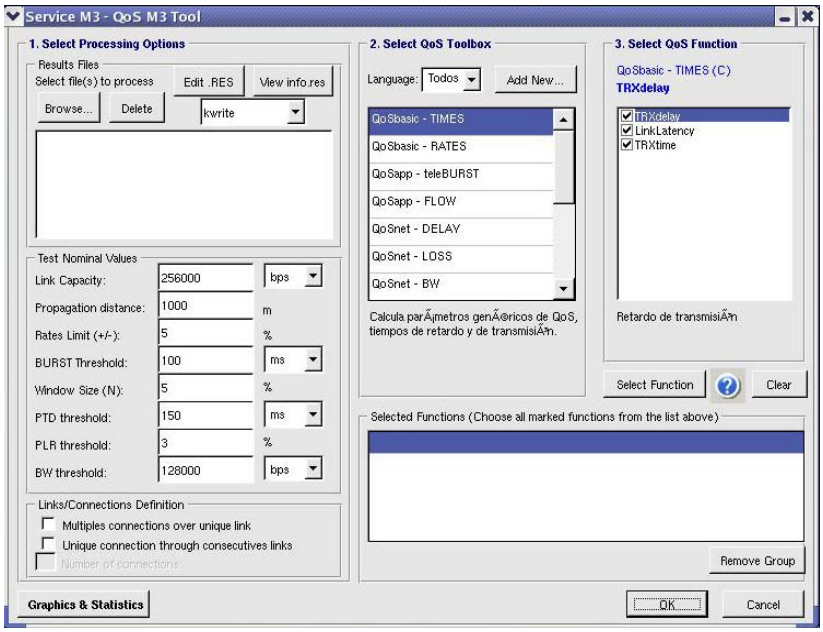


Fig. 1. Captura gráfica de la herramienta QoSM3.

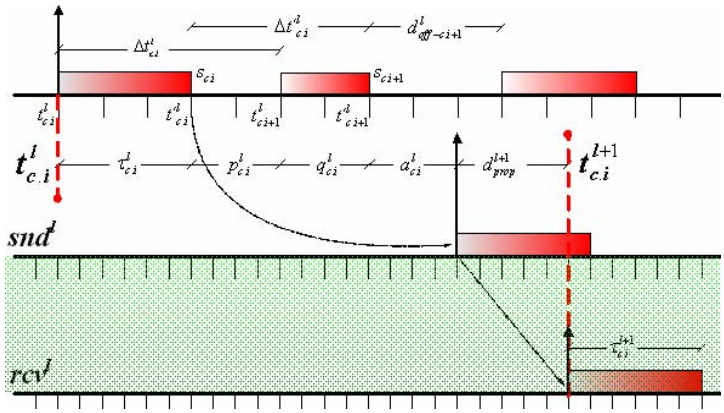


Fig. 2. Parámetros usados en el estudio (ver Apéndice I).

B. Módulo de flujo.

Este módulo aporta valor añadido al análisis estadístico realizado en el anterior. Calcula una serie de parámetros relacionados con el dimensionado de los *buffers* intermedios: tamaño total, ocupación instantánea, ocupación media, *throughput* en función de la capacidad nominal (C) del canal, factor de utilización ρ^* en (9), etc. Este factor ρ^* mide la relación entre el caudal real de datos y la carga media total ofrecida (que engloba el número de nodos, el tiempo medio entre llegadas y el tamaño de las tramas, conjuntamente), que coincidiría con el caudal eficaz (C_e) en un sistema sin pérdidas. Este valor siempre será, por tanto, menor que el máximo, $Ce_{máx}$, y por ello se normaliza cumpliendo que $\rho^* < 1$.

2.3 QoSM3 Red

La percepción subjetiva de la calidad de la transmisión suele ser valorada generalmente según tres parámetros: retardo, pérdidas y ancho de banda. Este bloque calcula dichos parámetros y los referencia a los umbrales de QoS asumibles en telemedicina. Está dividido en tres módulos:

A. Módulo de retardo.

Este módulo estima diferentes retardos según las condiciones de las medidas o simulaciones (entrada y salida de los *buffers* intermedios, de los enlaces, en los extremos de red, etc). Implementa los parámetros:

- *End-to-End Delay* (EED). Retardo de cada paquete, asumiendo los retardos necesarios indicados en QoSM3 Básico. Los umbrales, según la *Unión Internacional Telecomunicaciones* (UIT) [21] son:
  - 150ms, para vídeo con *codecs* H.261 y H.263.
  - 150ms, para audio con *codecs* G.711 (*Digital Subscriber Line*, DSL) y MRxx (*Global System Mobile*, GSM, y *Universal Mobile Transfer System*, UMTS).
  - 75ms, para *Voice over Internet Protocol* (VoIP) con *codecs* G.723.1, G.729, G.729A (sobre LAN).
  - 25ms, para telefonía clásica (sobre *Plain Old Telephony System*, POTS).
- *Max-EED*. Representa el *(1-α) cuantil* de la PDF del EED, donde se pierden las celdas que exceden el máximo. La red ajusta α para pérdidas asumibles.
- *Delay Variation (DV* o *jitter).* Varianza del retardo como diferencia entre retardos consecutivos. Para aplicaciones interactivas es necesario garantizar una probabilidad tal que P[*jitter* > 20ms] < 10%.

B. Módulo de pérdidas.
Este bloque cuantifica el nivel de pérdidas del sistema y caracteriza su comportamiento. Implementa el cálculo de *Packet Loss Rate (PLR)* como tasa de paquetes perdidos y muestra su evolución temporal y asociada a cada paquete transmitido. El diseño está implementado sobre aplicaciones basadas en *User Datagram Protocol* (UDP), mayoritarias en los servicios RT analizados en telemedicina. La inclusión de técnicas fiables de medida sobre *Transfer Control Protocol* (TCP) para el estudio de control de flujo y control de errores está en fase de desarrollo. Aun en estos casos, la tasa de perdidas sirve de indicador de congestión, observando la influencia de las retransmisiones en la comunicación. Para el caso de escenarios simulados, sin embargo, sí se encuentra operativo el procesado. A nivel teórico, igual que en el caso anterior, los umbrales de QoS asumibles para la PLR propuestos por la UIT [21] son:

- 0.10, para tecnologías FR y ATM.
- 0.05, para tecnologías ADSL.
- 0.03, para tecnologías UMTS.
- 0.01, para tecnologías LAN.

Además, este bloque también calcula la relación PLR vs EED, que permite valorar los servicios desde una perspectiva global. Esto ayuda a evaluar las prestaciones globales de QoS ofrecidas por el sistema completo, no sólo en cuanto a pérdidas o a retardo sino a la combinación de ambos parámetros.

C. Módulo de ancho de banda.
Este bloque completa la caracterización de QoS ya que es uno de los descriptores básicos de cualquier sistema de comunicación. Calcula parámetros determinantes para identificar los enlaces más restrictivos como aquellos que presentan menores recursos y, por tanto, limitan la capacidad global de la transmisión. Este ancho de banda máximo (*BandWidth*, BW) se refiere a los recursos de capacidad utilizables en un momento dado por las distintas aplicaciones que comparten un enlace. A partir de estos indicadores, detecta los cuellos de botella (*bottleneck*) como aquellos que acotan el BW disponible (*available BandWidth*, *aBW*) y que van a marcar, inevitablemente, los mínimos de capacidad para cada conexión. También evalúa el reparto de BW entre cada una de las aplicaciones para identificar los tráficos de no interés (*noise-traffic*) y el tráfico cruzado (*cross-traffic*).

Finalmente, este bloque implementa técnicas de cálculo instantáneo, acumulado y promediado para observar su evolución. Además, se incluye un algoritmo de estimación adaptativo, BE en (10), utilizando como factor de ponderación el retardo medido (si es extremo a extremo, EED; en general, indicando τ_e) según una ventana temporal (Δt_k) [22]. Con esta medida se obtiene una cuantificación muy interesante dependiente del retardo que contribuye, al igual que en la comparativa del bloque anterior (PLR vs EED) a una valoración global de la QoS.

3 Ejemplo de evaluación de servicio

Para el análisis de QoS se ha seguido una metodología de evaluación técnica presentada en un trabajo anterior [17] sobre el escenario planteado en Fig. 3, que incluye las siguientes características:

- **Modelo de aplicación.** Con dos tipos de servicios:
  - *Real Time*, RT, basados en *codecs Adaptative MultiRate* (AMR) para audio a 12.2Kbps, H.263 para vídeo uniforme a 16Kbps, y *RT Transport Protocol* (RTP) para transmisiones de señales electrocardiográficas (ECG) a 5Kbps.
  - *Store&Forward*, SF, basados en sesiones Telnet y Web, y transmisión *File Transfer Protocol* (FTP).
- **Modelo de red de acceso:** Ethernet conmutada, correspondiente a un acceso de *Local Area Network* (LAN) intra-hospitalaria que multiplexa un número variable de conexiones RT y SF.
- **Modelo de red troncal:** Basada en tecnologías *Frame Relay* (FR) y *Asynchronous Transfer Mode* (ATM), correspondientes a conexiones *Wide Area Network* (WAN) inter-hospitalarias.

El proceso de evaluación, como se ha explicado en la sección 2, se basa en un fichero de trazas obtenidas tanto en un entorno real de pruebas de laboratorio, como en su equivalente de simulación. El planteamiento teórico, según el escenario de red propuesto, considera una serie de datos por defecto incluidos en el Apéndice I: número de enlaces (L), distancia de propagación, (D_{prop}), y tiempo de bit (d_{bit}) como inversa de la capacidad nominal de transmisión (C). A partir de ellos y de las trazas de tráfico, se genera una estructura de datos formada por cuatro parámetros (definidos para cada conexión *c-ésima*, medidos en cada enlace *l-ésimo*, e identificados para cada paquete *i-ésimo* como $x^l_{c.i}$): marca temporal *timestamp* ($t^l_{c.i}$), tiempo entre paquetes ($\Delta t^l_{c.i}$), tiempo de OFF ($d_{off\text{-}i}$) y tamaño de paquete ($s_{c.i}$). Esta información es la base sobre la que se calcula el resto de parámetros que permiten obtener múltiples indicadores de teletráfico, relaciones temporales, etc. El esquema completo del diagrama de tiempos se ha mostrado en Fig.2.

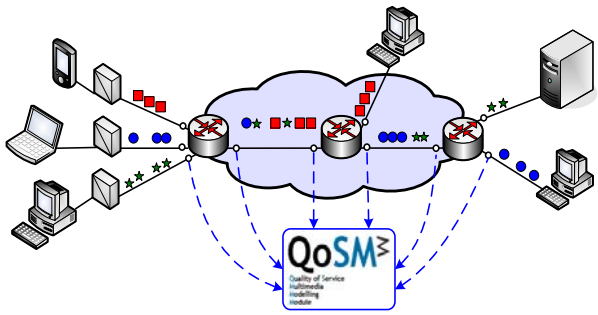


Fig. 3. Escenario genérico de evaluación.

3.1 QoSM3 Básico.

Este primer bloque se ha utilizado para caracterizar los aspectos básicos del servicio RT de audio basado en el *codec* AMR en un entorno real de telemonitorización hospitalaria [23]. La naturaleza RT de esta aplicación se podría inferir a partir de la tasa instantánea, de la tasa sostenida y del retardo de transmisión, representados en Fig. 4(a), 4(b) y 4(c), respectivamente. La evaluación conjunta de todos los casos muestra valores de $\Delta t^{l}_{c.i}$ y $s_{c.i}$ que pueden considerarse de escasa variabilidad. Además, calculando la PDF obtenida en el módulo siguiente (QoSM3 Aplicación) se puede concluir que la existencia de sólo dos ocurrencias significativas en $\Delta t^{l}_{c.i}$, corresponden a un proceso ON-OFF, ver Fig. 5. Todo ello permitiría caracterizar el servicio como tráfico de tasa constante (*Constant Bit Rate*, CBR) a 12.4Kbps según un modelo ON-OFF no homogéneo.

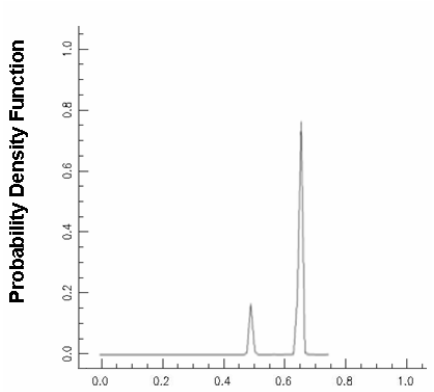
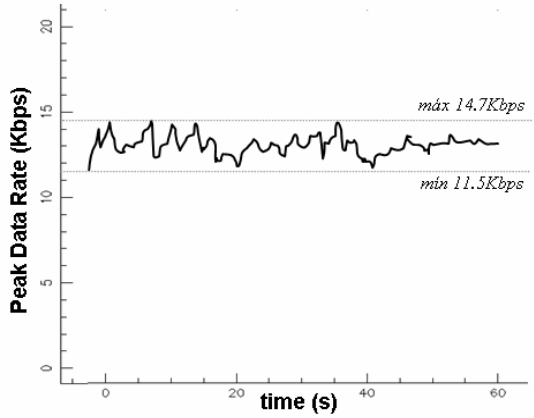
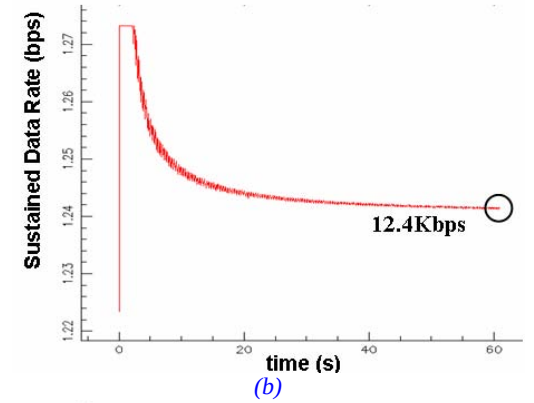


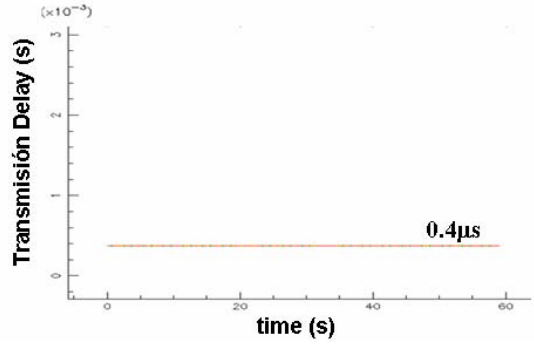
Fig. 5. PDF asociada al servicio RT de audio.



(a)

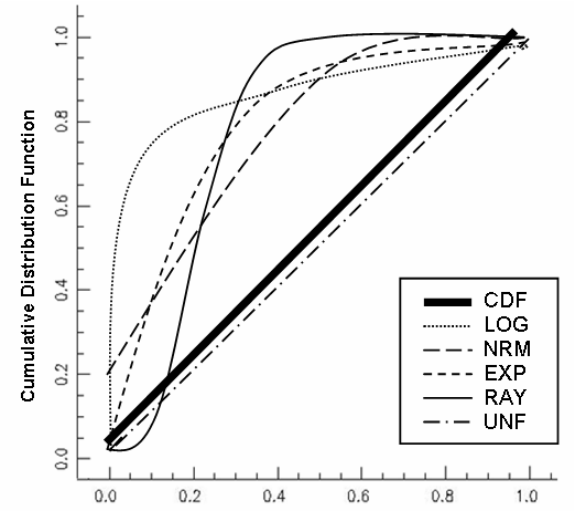


(b)

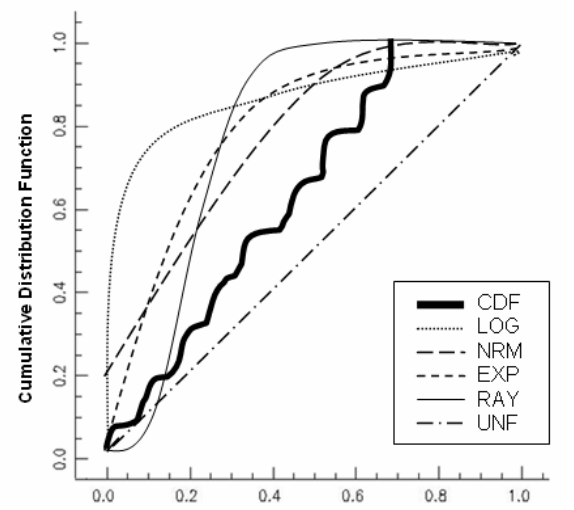


(c)

Fig. 4. QoSM3 Básico aplicado a servicio RT de audio. (a) PDR (bps) (b) SDR (bps) (c) $\tau^{l}_{c.i}$ (s)



(a)



(b)

Fig. 6. Comparativa de la CDF medida respecto a las distribuciones clásicas: log-normal, normal, exponencial, Rayleigh, y uniforme para: (a) servicio RT de video (b) Netmeeting

3.2 QoSM3 Aplicación.

A partir del módulo de ráfagas implementado en este bloque se puede llevar a cabo el reconocimiento y validación de modelos de tráfico. Como se ha citado, del análisis estadístico de $\Delta t^{l}_{c.i}$ y $s_{c.i}$, pueden deducirse las características y distribución del tráfico observado, así como equipararlas tanto a las tendencias clásicas del modelado (distribuciones exponencial, Pareto, Rayleigh, normal, log-normal, uniforme, etc.) como a las trazas o réplicas del mismo tráfico en diferentes puntos de observación. Esta idea se plasma para dos servicios RT de vídeo basados en el *codec* H.263 y en el *software* comercial *NetMeeting*, como muestra Fig. 6. Para el primer caso, se aprecia cómo el comportamiento sigue una distribución uniforme, propia del *codec*. Para cuantificar estas tendencias, QoSM3 ofrece la posibilidad de realizar una comparación númerica entre distribuciones estadísticas mediante el test de Kolmogorov-Smirnov (K-S) [24]. Dicho test K-S obtiene la distancia entre distribuciones, a partir de la cual se cuantifica y valora la tendencia que siguen las medidas. En Fig. 7 se muestran los datos calculados para el servicio RT de vídeo H.263, comparando sólo entre muestras significativas del tráfico observado, dónde se justifica su tendencia uniforme.

En el segundo caso, de la comparativa de CDF clásica mostrada en Fig. 6 no se concluye una característica evidente, por lo que habría que analizar los estadísticos de orden superior. Así, el rango de dependencia temporal puede evaluarse con la función de autocorrelación (ver Fig.8) y el cálculo del grado de subexponencialidad y del parámetro de Hurst (ver Fig.9). Dichos indicadores se obtienen mediante el estimador de Hill [18] y el método de la varianza de los residuales propuesto por [20], respectivamente. De ambos resultados, estimados para el servicio RT de vídeo *NetMeeting*, se justifica que la variabilidad observada podría seguir una tendencia autosimilar, que posiblemente es debida a la redundancia que genera el *codec* tanto temporal como espacialmente.

Finalmente, este módulo completa al bloque básico, como muestra Fig. 10, donde se aprecian dos valores significativos del tamaño instantáneo de ráfaga para el servicio de audio RT. Esto vuelve a justificar el comportamiento ON-OFF evaluado en Fig. 4 y 5.

```
Res-soloVideoFPS15_burst.log
Test K-S (972 muestras comparadas)
-------------------------------------------
Distribucion Teorica:  expo
Error Med= 0.176813    Var Error = 0.009216
Error MAX= 0.295765         en X = 0.368000
-------------------------------------------
Distribucion Teorica:  norm
Error Med= 0.133121    Var Error = 0.002470
Error MAX= 0.187890         en X = 0.001000
-------------------------------------------
Distribucion Teorica:  logn
Error Med= 0.286837    Var Error = 0.031622
Error MAX= 0.532270         en X = 0.094000
-------------------------------------------
Distribucion Teorica:  rayl
Error Med= 0.174757    Var Error = 0.014007
Error MAX= 0.345548         en X = 0.479000
```

Fig. 7. Test de Kolmogorov-Smirnov.

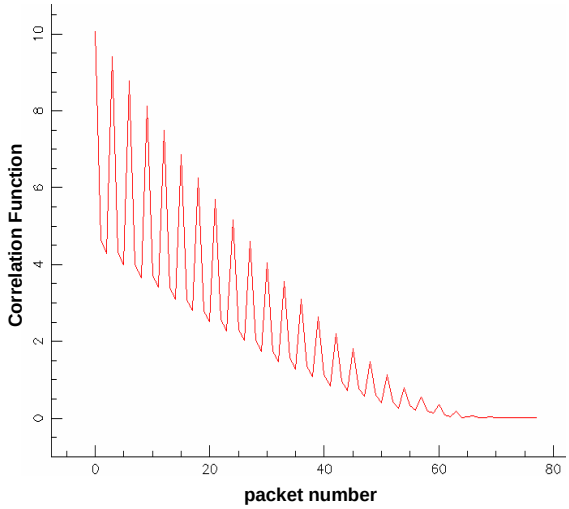


Fig. 8. Función de autocorrelación para el servicio RT de vídeo Netmeeting.

```
Res-influRecEnvioVideoYAudio
Subexponencialidad y Fractalidad:
-----------------------------------------
estimador de Hill           alfa = 0.594672
Varianza de los Residuales     H = 0.820632
```

Fig. 9. Grado de subexponencialidad y parámetro de Hurst.

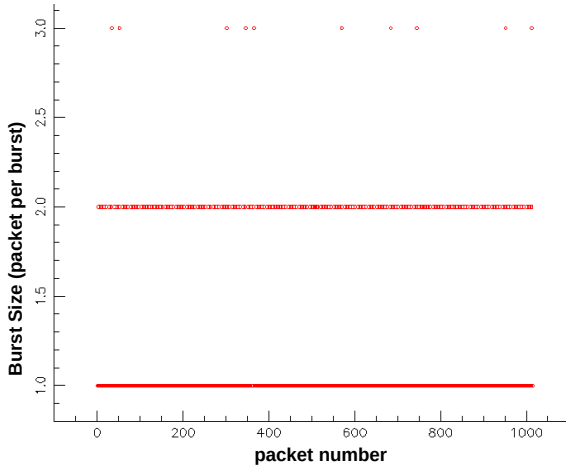
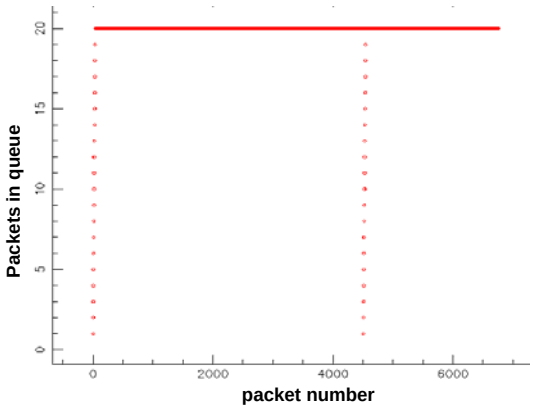
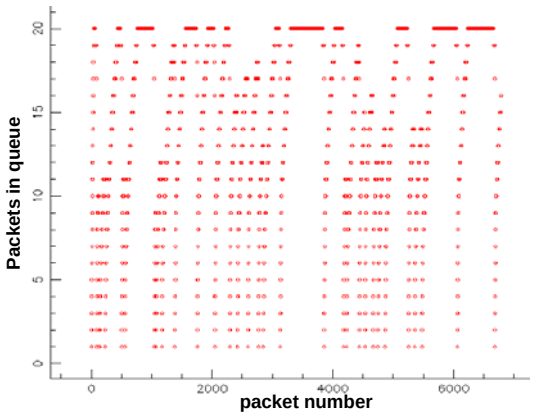


Fig. 10: Tamaño instantáneo de la ráfaga.

Como se introdujo en la sección 3, este bloque incluye un módulo de flujo que complementa el modelado anterior del servicio multimedia, con una caracterización de su comportamiento en un entorno de red, incluyendo el análisis de los *buffers* intermedios. Los resultados analizados corresponden a un servicio SF basado en transmisiones FTP y medido en los distintos puntos intermedios de un escenario simulado como el planteado en sección 2. Un caso representativo se muestra en Fig. 11 que indica el nivel medio de ocupación para capturas en estados opuestos de carga de la red. Así, Fig.11(a) corresponde a una única conexión que no congestiona el enlace y permite una ocupación prácticamente constante ($\rho^*\approx 1$) del *buffer*, de tamaño Q=20 paquetes. Fig. 11(b) corresponde a la multiplexación de tres conexiones, que sí congestionan el enlace, y hacen fluctuar el factor ρ^*.

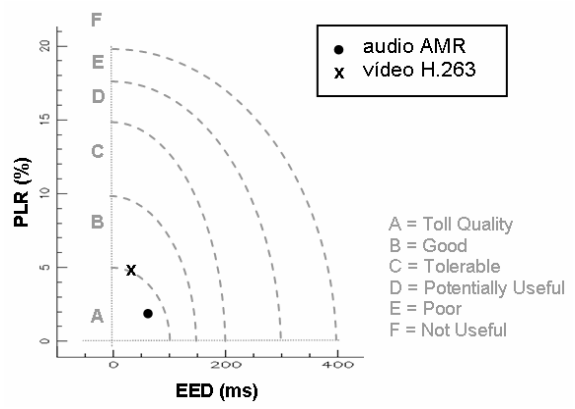


(a)

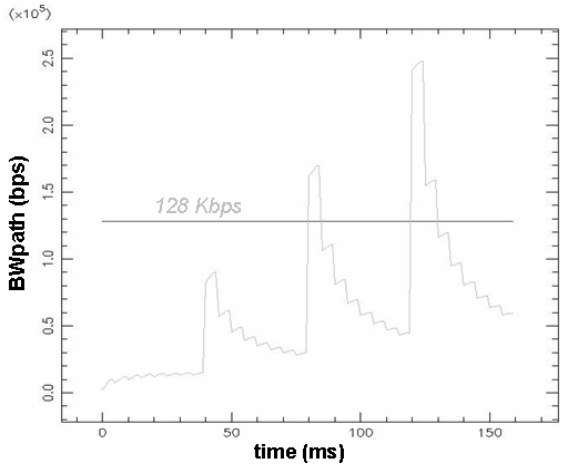


(b)

Fig. 11. Ocupación instantánea del buffer. (a) Sin congestión. (b) Con congestión.



(a)



(b)

Fig. 12. QoSM3 Red para servicios RT. (a)PLR vs EED. (b) Evolución de BW sobre un path de 4 enlaces.

3.3 QoSM3 Red.

Los parámetros básicos que evalúa este último bloque son las pérdidas (PLR), el retardo (EED) y el ancho de banda (BW), junto con sus cálculos asociados. Se muestran los datos obtenidos para los servicios RT tanto de audio AMR como de vídeo H.263. Además de los parámetros clásicos, explicados en la sección 2, se han implementado diversas técnicas para estimar el BW a partir de sus valores instantáneos, de promediado basado en ventanas configurables, y de cálculo adaptativo con el estimador definido en (10). Este estimador aporta tendencias muy interesantes ya que utiliza factores de ponderación dependientes del retardo, calculado como estimación de BW de banda extremo a extremo. Esta doble estimación hace más fiables los valores y permite anticiparse a las posibles situaciones de congestión que se den en la red.

Siguiendo esta línea, la herramienta QoSM3 también incluye la relación entre pérdidas y retardo (PLR/EED), como muestra Fig. 12(a). Este ratio permite evaluar si el servicio estudiado cumple los requisitos de QoS: es decir, si se localiza dentro de las posibles áreas de utilización, marcadas en gris y tomadas de [25]. En el caso presentado, se aprecia cómo ambos servicios RT cumplen perfectamente los umbrales, si bien el de audio presenta mejores indicadores que el de vídeo.

Así, mediante esta inspección visual es directo comprobar si el tráfico observado está dentro de los márgenes permitidos o no, en cuyo caso, se podría actuar en consecuencia modificando los parámetros de generación. Este planteamiento es muy interesante ya que permitiría evitar las situaciones de congestión por anticipado y de forma adaptativa sin necesidad de modificar las configuraciones de los elementos intermedios de la topología, sino variando los parámetros de las aplicaciones generadoras de tráfico (tasas, *codecs*, ratios de compresión, etc.).

Por último, señalar que estas medidas se pueden realizar sobre una conexión (*none*), varias conexiones en un enlace (*link*), una conexión a lo largo de enlaces consecutivos (*path*), o varias conexiones en un *path* (*cross*). Ello permite detectar enlaces restrictivos, cuellos de botella, o instantes en los que BW rebasa el umbral de QoS y se deben corregir los parámetros de generación (tasas, *codecs*, prioridades, etc.) para evitar esas situaciones. Así, se muestra en Fig. 12(b) un ejemplo de evolución del BW conjunto generado por el servicio RT de audio y vídeo, a lo largo de cuatro enlaces. Como puede observarse, se comprueba de forma directa que el tráfico sobrepasa el BW disponible de 128Kbps en los enlaces 3 y 4.

4 Conclusiones

En este trabajo se plantea una herramienta que, siguiendo una metodología de evaluación técnica presentada en trabajos anteriores y, sin perder generalidad, permite realizar un estudio específico de la QoS que ofrece la red y del modelo de tráfico tratado. Esto permite evaluar el diseño de las fuentes generadoras de datos y, por tanto, de las aplicaciones que influyen en el sistema completo a analizar. En el estudio, concretamente, se ha utilizado la herramienta para caracterizar servicios multimedia de audio y vídeo a tiempo real, en entornos de telemedicina

El modelo de tomografía de red implementado en la herramienta QoSM3, permite realizar un análisis cuantitativo y cualitativo de las distintas partes de un escenario propuesto. A partir de la interpretación de las medidas, se puede determinar características intrínsecas al servicio y al medio sobre el que se transmite (QoS básica), modelar el tráfico generado, transmitido o de entrada y salida a un *buffer* (QoS de aplicación), y caracterizar el enlace o dispositivo de comunicación sobre el que se mide (QoS de red).

Apéndice I. Variables utilizadas

L	**Número de enlaces** (*path length*) que constituyen el camino extremo a extremo con diversas tecnologías.
D^l_{prop}	**Distancia de propagación** (*m*).
d^l_{bit}	**Tiempo de bit** (*s*). Duración de cada bit en el enlace *l-ésimo*. Su inversa es la capacidad $c^l = 1 / d^l_{bit}$ (*bps*)
$t^l_{c.i}$ $t'^l_{c.i}$	**Marca temporal - timestamp** (*s*). Instante temporal en que el *primer bit* de cada paquete *i-ésimo* de la conexión *c-ésima* del enlace *l-ésimo* se transmite. Su equivalente para el *último bit*, se marca con ´.
$\Delta t^l_{c.i}$ $\Delta t'^l_{c.i}$	**Tiempo entre paquetes** (*s*). Intervalo temporal entre el primer bit de dos paquetes consecutivos (*i, i+1*). Su equivalente para el último bit, se marca con ´.
$d^l_{off-c.i}$	**Tiempo OFF** (*s*). Intervalo temporal entre el *último bit* del paquete *i-ésimo* y el *primer bit* del *i+1-ésimo*.
$s_{c.i}$	**Tamaño de paquete** (*bytes*).
$\tau^l_{c.i}$	**Retardo de transmisión** (*s*). Intervalo temporal desde el *primer bit* hasta el *último bit* que accede al medio, según $\tau^l_{c.i} = t'^l_{c.i} - t^l_{c.i} = s_{c.i} \cdot d^l_{bit}$
$p^l_{c.i}$	**Retardo de procesado** (*s*). Incluye una parte variable (routing, colas..) y otra fija (encapsulado, fragment..)
$q^l_{c.i}$	**Retardo en cola** (*s*). Tiempo de espera en cola según la disciplina, nº paquetes/conex, nº conexiones, etc.
$a^l_{c.i}$	**Retardo de acceso al medio** (*s*). Tiempo que espera cada paquete a acceder al siguiente medio. Depende de la técnica de acceso, tecnología, nº conex., etc.
d^l_{prop}	**Retardo de propagación** (*s*). Intervalo temporal desde el *primer bit* que accede al medio hasta el *primer bit* que se recibe; es decir, la *latencia* de enlace.

Agradecimientos

Este trabajo ha recibido el apoyo de proyectos de la Comisión Interministerial de Ciencia y Tecnología (CICYT) y de los Fondos Europeos de Desarrollo Regional (FEDER) TSI2004-04940-C02-01, de los Fondos de Investigación Sanitaria (FIS) FISG03/117, y del Ministerio de Educación y Ciencia (beca FPU AP-2004-3568).

Referencias

[1] G. Fortino and L. Nigro, "A methodology centered on modularization of QoS constraints for the development and performance evaluation of multimedia systems", *Proc. 33rd Annual Simulation Symposium SS'00*, pp. 177-184, 2000.

[2] R. Holle and G. Zahlmann, "Evaluation of telemedical services", *IEEE Trans Inf Technol Biomed*, 3(2):84-91, 1999.

[3] S. de Lusignan, S. Wells, P. Johnson, K. Meredith, E. Leatham "Compliance and effectiveness of 1 year's home telemonitoring. The report of a pilot study of patients with chronic heart failure". *Eur J Heart Fail*, 3(6):723-30, 2001.

[4] M. Maheu, P. Whitten, A. Allen, "E-health, telehealth and telemedicine: guide to start-up &success", *Jossey-Bass*, 2001.

[5] W.C. Hardey, "QoS Measurement and Evaluation of Telecommunications Quality of Service", *Eds. John Wiley*, Handcover, 230 pages. [Book Review – R. Chodoreck, *IEEE Communications Magazine*, 40(2):30-32, 2002].

[6] K. Zielinsky, "Krakow Centre of Telemedicine – Developing the Platform for Regional Telemedical Networks", *Proc. Conference 'E-health in Common Europe'*, 2003.

[7] W.R. McDermott et al., "Optimization of Wide-Area ATM and Local-Area Ethernet/FDDI Network configurations for high-speed telemedicine communications employing NASA's ACTS", *IEEE Network*, 13(4):30-38, 1999.

[8] D. Caramella and S. Giordano "An advanced IP based telemedicine trial supporting quality of service for multimedia teleconsulting", *International Conference EuroPACS*, 2000.

[9] J.F. Huard, I. Inoue, A. A. Lazar and H. Yamanaka, "Meeting QoS guarantees by end-to-end QoS monitoring and adaptation", *V IEEE International Symposium on High Performance Distributed Computing*, pp. 348-355, 1996.

[10] I. Martínez, J. Salvador, J. Fernández, J. García, "Traffic requirements evaluation for a Telemedicine network", *International Congress on Computational Bioengineering ICCB'03*, pp. 389-394, 2003.

[11] M. Li and R. Sampigethaya. "Network Tomography", *STAT 593E*, 2003.

[12] T. Bu, N.G. Duffield, F. Lo Presti and D. Towsley. "Network Tomography on General Topologies". *Umass CMPSCI Technique* Report, 2002.

[13] X. Fang and D. Ghosal, "Performance modelling and QoS evaluation of MAC/RLC layer in GSM/GPRS Networks", *IEEE International Conference on Communications ICC'03*, vol. 1, pp. 271-275, 2003.

[14] P.E. Wirth, "The role of teletraffic modeling in the new communications paradigms," *IEEE Communications Magazine*, vol. 35, pp. 86-92, 1997.

[15] A. Vogel, G. Bochmann, R. Disallow, J. Geckos and B. Kerherv, "Distributed Multimedia Applications and Quality of Service – A survey", IEEE Multimedia, 1994.

[16] W. Willinger and V. Paxon, "Discussion of 'heavy tail modelling and teletraffic data' by S.R. Resnick", The Annals of Statistics, 25(5):1856-1865, 1997.

[17] I. Martínez J. García and J. Fernández. "QoS Evaluation Methodology for Multimedia Telemedicine Services". *IEEE Transactions on Multimedia*, submitted, 2005.

[18] E. Casilari. "Caracterización y modelado de tráfico de vídeo VBR", *Tesis Doctoral*, Universidad de Málaga, 1998.

[19] A. Coppola, "Practical Statistical Tools for Reliability Engineers", *DoD Reliability Analysis Center*, 1999.

[20] M. Taqqu, "Theory and Applications of Long-Range Dependence", *Birkhäuser*, Boston, 2003.

[21] D. Wright, "Informe UIT sobre telemedicina en los países en desarrollo", *Journal of Telemedicine and Telecare*, 4(1), 1998 [Versión española en *International Telemedicine*,7-8, 1998].

[22] M. Valla, R. Wang, M. Gerla and M.Y. Sanadidi. "TCP Westwood, Bandwidth estimation techniques for efficient and friendly congestion control", University of California, *http://www.cs.ucla.edu/NRL/hpi/tcpw/*, Last access 30/03/05.

[23] E. Viruete, C. Hernández, J. Ruiz, J. Fernández, A. Alesanco, E. Lleida, A. Ortega, A. Hernández, A. Valdovinos, J. García, "Sistema de telemonitorización en vehículos de emergencias médicas sobre UMTS", *Proc. CASEIB*, pp. 111-114, 2004.

[24] J.L. Romeu, "K-S: A goodness of fit test for small samples", *START Reliability Analysis Center*, vol. 10, number 6, 2003.

[25] T.J. Kostas et al. "Real-Time voice over packet-switched networks", *IEEE Networks*, 12 (1): 18-27, 1998.

Estudio de un Router Software para la implementación de una Pasarela Residencial

Jaime García, Francisco Valera, David Díez, Hugo Gascón, Carmen Guerrero, Arturo Azcorra
Departamento de Ingeniería Telemática. Universidad Carlos III de Madrid
Avda. de la Universidad, 30.
28911 - Leganés (Madrid)
E-mail: jgr, fvalera, ddiez, hgascon, guerrero, azcorra@it.uc3m.es

Abstract *Residential Gateways (RGW) are the last (or first, it depends on the point of view) element in the access network and the interface between Home Network and access provider. In MUSE [1] our work is related to the design of a RGW compliant with QoS requirements and to the development of a RGW prototype. In this paper we propose a new model to design RGW, based on the Click modular router [2], but using user level applications too. This is what we call an Hybrid Model due to the capacity to create new functionalities at different levels: as a normal application or extending Click. This capacity helps the developer to extend the RGW device. A complete set of tests validations are presented to support this innovative idea.*

1. Introducción

El proyecto europeo MUSE [1] tiene como misión la especificación de la arquitectura y el modelo de la futura red de datos de banda ancha para Europa, con la idea de crear un acceso de bajo coste y alta calidad para los usuarios. Aunque las redes de acceso y central son unos puntos importantes de investigación, los dispositivos ubicados en los entornos residenciales son también componentes clave para poder proporcionar buenas calidades de servicio, especialmente los llamados Residential Gateway (RGW) o Pasarelas Residenciales.

La RGW es el primer elemento de la red accesible por el usuario como se muestra en la Fig. 1. Cada dispositivo del hogar estará conectado (de forma directa o inalámbrica) a través del RGW a una red de banda ancha, pero compartida. Por lo tanto, información de tiempo real como pueden ser datos de voz o video pueden estar compartiendo el mismo medio que otras conexiones menos prioritarias como conexiones web. Como se puede ver, algún tipo de *priorización* es necesaria antes de que los paquetes salgan a la red. Además, se necesita conformar el tráfico de los distintos flujos para administrar el ancho de banda disponible.

Además de todas estas funcionalidades, la RGW debe realizar muchas otras tareas por el usuario final: auto configuración, control, administración, configuración de los servicios, etc. Existen algunas funcionalidades que no se describirán en este artículo, como la señalización de la calidad de servicio (QoS). En MUSE se están estudiando diferentes modelos para la señalización y provisión de la QoS extremo a extremo, centrándose en la red de acceso. Una de las alternativas consideradas es IP Multimedia Subsystem (IMS) [3] para la configuración de los servicios. Realmente esta especificación tendrá que ser adaptada al escenario de redes de acceso fijas ya que hoy en día el IMS está sólo especificado para escenarios móviles y de redes WLAN. MUSE trabajará junto al ETSI-TISPAN [4] para realizar esta adaptación, en la que la RGW tiene un papel importante. El prototipo de la RGW en MUSE incorporará esta funcionalidad, no en las primeras etapas del desarrollo, pues el trabajo se tiene que hacer en paralelo con la estandarización del ETSI-TISPAN.

Se ha escogido utilizar el sistema operativo Linux para esta etapa de implementación del prototipo de la RGW que se ejecutará en un dispositivo PC compatible *i386* de reducidas dimensiones [5]. Debido a que la RGW debe manejar paquetes de bajo nivel (del nivel de enlace) y Linux no provee mecanismos para realizar esta manipulación de forma nativa, se ha decidido utilizar el Click modular router [2] para el tratamiento de las tramas de nivel de enlace. Es importante señalar que los resultados que aquí se presentan son extrapolables a otra herramienta software que trabaje al mismo nivel de Click, es decir, cualquier herramienta que capture tramas a nivel de enlace y sea capaz de enviarlas sin modificar al nivel de aplicación. En la sección 2 se hará una introducción a las principales características de Click así como a su elección como potencial herramienta de desarrollo. La sección 3 se dedica a la descripción del modelo propuesto para la implementación de la RGW. La sección 4 describe los escenarios de pruebas planeados y las pruebas ya realizadas. Por último, la sección 5 finaliza con las principales conclusiones extraídas de todas las pruebas realizadas.

[1] MUSE (Multi Service Access Everywhere) es un gran proyecto de I+D en las redes de banda ancha. Dentro del sexto Programa Marco, MUSE contribuye al objetivo estratégico de "Ancho de Banda para Todos" dentro de las IST (Information Society Technologies) y está parcialmente financiado por la Comisión Europea.

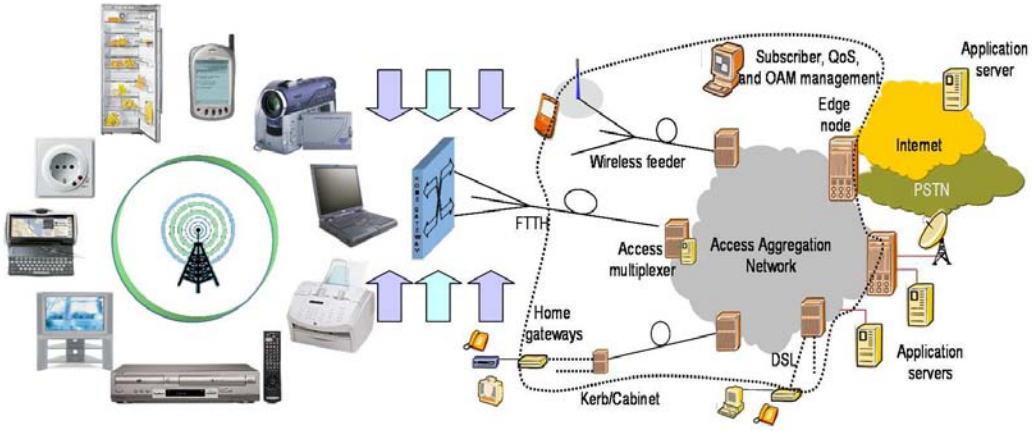


Figura 1: La red de banda ancha en MUSE

2. Plataforma Click

Click [6] es un *software router* modular desarrollado conjuntamente por el grupo LCS's Parallel and Distributed Operating Systems del MIT, Mazu Networks, el ICSI Center for Internet Research y últimamente UCLA. Un kernel de Linux ejecutando Click es capaz de actuar como un router de una forma flexible y configurable. Las tareas de encaminamiento se realizan de una manera extremadamente rápida para un software ejecutándose en un hardware común. En un Pentium III a 700 MHz, un router IP con Click puede manejar hasta 333.000 paquetes de 64 bytes por segundo [2].

Configurar un router con Click se basa en interconectar módulos llamados *elementos* que controlan cada aspecto de la operación del router: comunicación con los dispositivos, modificación de paquetes, colas, políticas de descarte, envío de paquetes, etc. Como la modularidad es la principal ventaja de Click, es posible escribir nuevos elementos en C++ con la funcionalidad requerida. La configuración del router resulta de la unión de elementos en un fichero de configuración usando un lenguaje simple propio de Click.

En Click se puede trabajar a nivel de *aplicación* o usando un *módulo del kernel* de Linux, cambiando ligeramente el funcionamiento del mismo. Se ha escogido la segunda opción para la realización del prototipo, ya que con ella se permite la manipulación de tramas de una manera más rápida, eliminando los retardos introducidos por la pila de protocolos TCP/IP. Actualmente Click sólo funciona en kernel de Linux de la versión 2.2 y 2.4 aunque se espera que pronto sea implementado para las versiones 2.6. Como es un proyecto de software abierto, muchos desarrolladores crean nuevas funcionalidades y varios están trabajando para soportar el kernel 2.6 y añadir nuevas funcionalidades para IPv6.

3. Modelo Híbrido Click / Aplicación

Para la implementación del prototipo de la RGW en el proyecto MUSE se necesita un software capaz de capturar todos los paquetes de la capa de nivel de enlace, modificarlos y volver a enviarlos a la red; enviarlos al nivel de aplicación, etc. Por lo tanto, se decidió utilizar Click (exactamente el módulo kernel de Click). Como hemos visto en el apartado 2, en Click se encuentran implementadas varias funcionalidades, pero en nuestro desarrollo necesitaremos crear nuevas. Existe la posibilidad de crear nuevos elementos de Click que se integren con los que actualmente existen. Otra posibilidad sería capturar tramas a nivel de Click y enviarlas todas al nivel de aplicación para ser procesadas ahí. Realizarlas de una forma u otra presenta los siguientes problemas:

1. Programar nuevos elementos a nivel de kernel es, en ocasiones, considerablemente complejo y más aún cuando se trata de aplicaciones con una gran carga de componentes de servicios de red.

2. La programación de nuevas aplicaciones hardware o software pueden ser necesarias en el futuro y deben crearse independientes de la plataforma cuando sea posible (desarrollándolas en Java, por ejemplo).

Sin embargo es evidente que trabajar a nivel de aplicación presenta además de problemas de eficiencia, problemas técnicos importantes cuando hay que gestionar los niveles más bajos de la torre de protocolos. Para solventar estos problemas, se decidió crear un nuevo *modelo híbrido* donde no se usa un modelo puro de Click o de aplicación, sino que se crean nuevas implementaciones de software en el nivel o capa más conveniente (como un elemento de Click o como una aplicación, según sea mejor). La Fig. 2 muestra el *modelo híbrido* presentando tres bloques principales:

Click es la implementación de Click trabajando a nivel de kernel. Recibirá cada paquete, comprobará el tipo de paquete, y lo enviará directamente por otros elementos de Click o lo enviará hacia el nivel de aplicación según tenga configurado.

El Manager recibirá paquetes *nuevos* del módulo Click y los procesará. Dependiendo de las características del paquete, el Manager puede configurar el módulo Click para que envíe paquetes con las mismas características a otro proceso de nivel de aplicación.

Procesos P1-Pn son las aplicaciones de nivel de aplicación desarrollados para realizar ciertas funciones.

Esta propuesta de modelo debe ser validada primero para comprobar que su rendimiento es aceptable para el funcionamiento de un RGW. Cuando se realizan todas las operaciones al nivel de Click nos ahorramos el tiempo de tránsito de un paquete por la pila TCP/IP del kernel de Linux. En el modelo híbrido propuesto, este tiempo de tránsito debe ser otra vez tenido en cuenta ya que algunos paquetes (paquetes de señalización por ejemplo) se enviarán desde el nivel Click al nivel de aplicación. Debido a esto es imprescindible estudiar el retardo impuesto por enviar una trama desde el nivel de Click al Manager puesto que este retardo limitará y marcará las posibilidades que tiene esta plataforma como base del desarrollo de la pasarela residencial para un entorno de acceso a banda ancha como el planteado en el proyecto MUSE.

4. Validación del Modelo Híbrido

4.1. Retardo de procesamiento de paquetes

Con esta prueba se va a medir el retardo adicional que supone enviar los paquetes desde el nivel

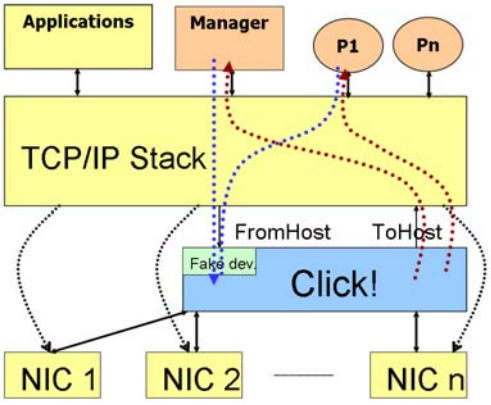


Figura 2: Modelo híbrido

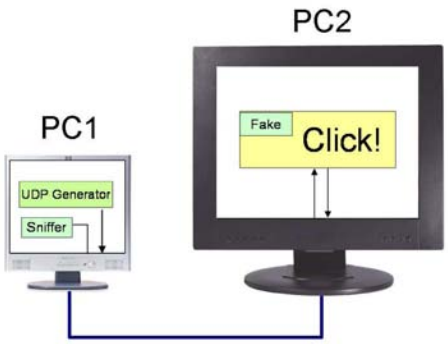


Figura 3: Conexión directa

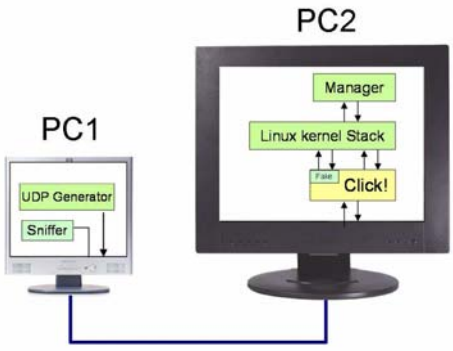


Figura 4: Conexión con Manager

Click hacia la aplicación Manager y de vuelta a Click.

Se crearon dos escenarios para comparar estos valores (Figs. 3 y 4). En ambos escenarios se utilizan dos ordenadores directamente conectados. Un ordenador actúa de generador y el otro ejecuta Click que, en cada escenario, realiza una función diferente. En el primer escenario denominado *conexión directa* el ordenador con Click cambia las cabeceras Ethernet e IP para devolver inmediatamente el paquete al ordenador origen. En el otro escenario (*conexión con Manager*) el nivel Click encapsula la trama entera en un paquete UDP y lo envía hacia la aplicación Manager que lo devuelve inmediatamente al nivel Click que realiza el mismo cambio de cabeceras que en el escenario anterior. En cada escenario se realizaron experimentos con diferentes tamaños de paquetes y se obtuvo la media de dichos experimentos. El resultado final se puede ver en la Tabla 1.

Como resultado de estas pruebas se puede concluir que la diferencia del retardo es independiente de la longitud del paquete y además que esta diferencia está tan sólo entre los 130 y 140μs. Hay que mencionar finalmente que estas pruebas se realizaron en un ordenador *compacto* con las siguientes especificaciones:

- Placa Lex Solution SV860A
- Procesador VIA C3 533MHz
- 512 Mbytes de RAM
- 2 tarjetas de red 10/100 BaseTx

Tamaño de paquete	Conexión directa	Conexión con Manager	Diferencia
100 bytes	$120\mu s$	$250\mu s$	$130\mu s$
540 bytes	$200\mu s$	$330\mu s$	$130\mu s$
1060 bytes	$290\mu s$	$430\mu s$	$140\mu s$
1440 bytes	$365\mu s$	$500\mu s$	$135\mu s$

Tabla 1: Retardo introducido por el modelo híbrido

- 1 tarjeta de red 10/100/1000 BaseTx
- 1 tarjeta de red inalámbrica Atmel 802.11b

Es importante destacar que *ToHost* fue el elemento de Click utilizado para las pruebas con el Manager. El elemento *ToHost* envía un paquete que recoge en su entrada hacia la pila TCP/IP de Linux. *ToHostSniffers* es otro elemento similar a *ToHost* que envía tramas hacia aplicaciones que leen directamente de los interfaces de red como haría la librería *libpcap* [7]. Debido a que en Java no existe un soporte nativo para leer directamente de un interfaz de red, se tuvieron que realizar pruebas con otras APIs que implementan una funcionalidad similar a la de *libpcap* para poder probar el elemento *ToHostSniffers* que disminuiría el tiempo de tránsito de un paquete al no tener que atravesar la pila TCP/IP de Linux. Las dos APIs probadas ([8] y [9]) se denominan *jpcap* y no mejoran los resultados dados por *ToHost*. Los resultados obtenidos se muestran en la Fig. 5. Se han eliminado los resultados de la API [8] ya que son mucho peores que los de [9].

Estos resultados pueden deberse a la forma de implementar las APIs de Java ya que hacen uso de la librería *libpcap* como paso intermedio. Por ello, para las siguientes pruebas sólo se programará el Manager para que trabaje con el elemento *ToHost*.

4.2. Carga soportada por el modelo

Con este escenario se desea probar si el uso de una aplicación en el nivel de aplicación, llamada el Manager, aumenta o no considerablemente el retardo de tránsito (si realmente se reduce el rendimiento utilizando un Manager se podría realizar las mismas funciones en el propio nivel de Click, pero esto reduciría la flexibilidad del desarrollo y complicaría el mismo). Para estas pruebas, se instaló Click en el ordenador *compacto* mencionado en apartados anteriores.

En el escenario de pruebas se conectaron dos ordenadores de diferentes redes IP a través de este ordenador que denominamos RGW. Este escenario se puede ver en la Fig. 6.

En estas pruebas se desea calcular el *throughput* y el *jitter* obtenidos en el equipo *servidor* al enviar paquetes desde el terminal *cliente*. Para comparar prestaciones, se configuró el equipo *RGW* para que realizase funciones de NAT (Network Address Translation). Se implementó el NAT de tres formas distintas para poderlas comparar:

1. Linux con IPTables. Para realizar esta prueba se necesita compilar el núcleo de Linux para incluir soporte para IP Tables. Una vez compilado, tan sólo es necesario activar el servicio de NAT con la instrucción *iptables -t nat -A POSTROUTING -o eth1 -j MASQUERADE*. También es necesario activar el *ip_forwarding* activando el bit del fichero */proc/sys/net/ipv4/ip_forward*.

2. NAT implementado en Click. En Click existen elementos que permiten realizar un NAT a este nivel.

3. NAT implementado en una aplicación (Manager). En estas pruebas el nivel Click envía todos los paquetes (utilizando el elemento *ToHost* de Click, como se comentó anteriormente) a una aplicación Java que realiza el NAT.

En las pruebas se ha utilizado el programa *iperf* [10] para la generación de tramas y obtención de estadísticas de las mismas. Existe un programa *cliente* que genera tramas a un ordenador donde se esté ejecutando una instancia *servidor* del programa *iperf*. Este programa *servidor* captura tramas y genera estadísticas del *throughput* y del *jitter*. La ventaja del *iperf* es que en el programa cliente podemos especificar la duración de la simulación, el tamaño de paquete y la tasa de generación de tramas.

Se realizaron las pruebas para los escenarios mencionados anteriormente y se utilizaron paquetes de 1470, 850 y 200 bytes para tener tamaños representativos. Los resultados se pueden ver en la Fig. 7. Se han omitido los resultados del escenario con IPTables ya que son idénticos a los del escenario con Click.

Los resultados más importantes que podemos extraer de estas pruebas son los siguientes:

- La máxima tasa de generación de paquetes depende del tamaño de dichos paquetes. En el caso de utilizar paquetes de 1470 bytes, la tasa máxima que puede generar el equipo *cliente* es

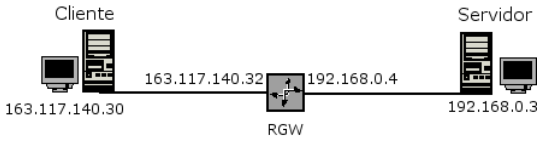


Figura 6: Primer escenario de pruebas

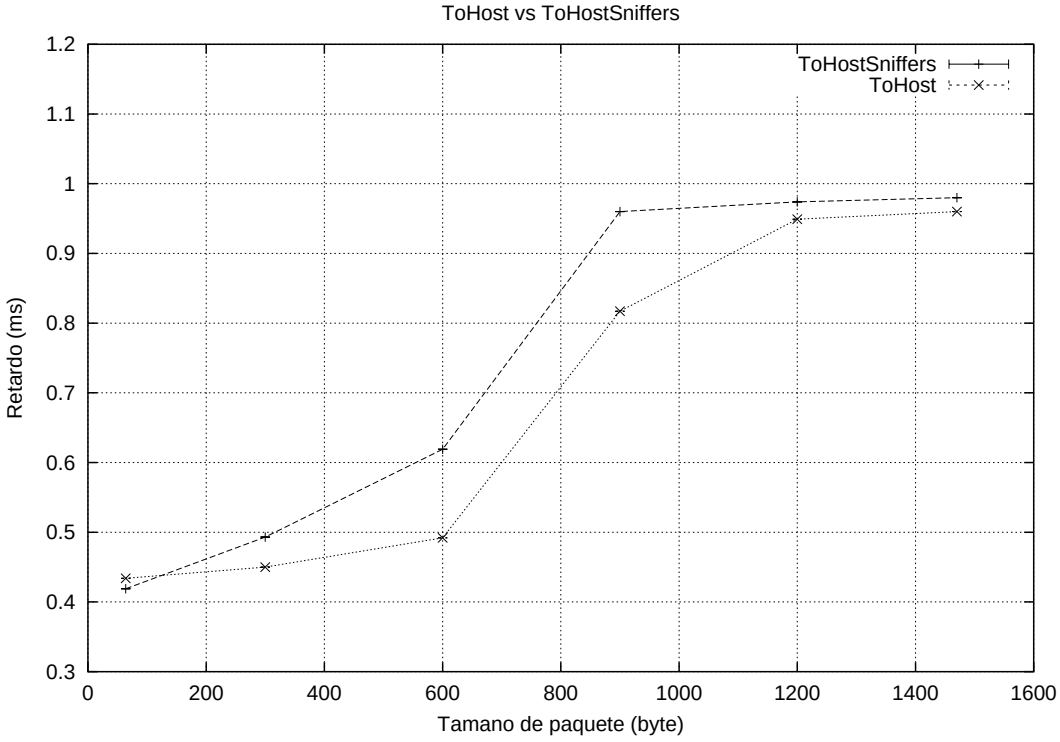


Figura 5: Resultados con la API jpcap

de 95,1 Mbps. Cuando se transmiten paquetes de 850 bytes ese máximo se encuentra en 92,7 Mbps mientras que para paquetes de 200 bytes no se pueden emitir más de 75,1 Mbps.

- Los escenarios de NAT con Click y NAT con IPTables dan resultados prácticamente idénticos.

- El escenario donde se realiza el NAT a nivel de aplicación (Manager) nos muestra resultados parecidos a los anteriores para paquetes mayores de 850 bytes y tasa de generación menor de 40 Mbps. Cuando nos salimos de estos valores el rendimiento ya no es tan bueno.

A la vista de estos resultados, es importante reseñar que estas pruebas están orientadas a comprobar el funcionamiento del modelo híbrido, donde algunas tramas pasarán del nivel Click al de aplicación. Se debe reafirmar el hecho de que los paquetes de datos atravesarán el nivel de Click sin "subir" al de aplicación por lo que sólo tramas de señalización y los primeros paquetes de flujos nuevos tendrán que ser procesados en dicho nivel.

Viendo los resultados mostrados en la Fig. 7 se puede ver que el tamaño de los paquetes de señalización, que han de subir al Manager y ser procesados de manera diferente a los paquetes de datos, influye en el throughput conseguido en el modelo híbrido. Si, como hemos indicado anteriormente, se considera señalización para la configuración de los servicios mediante el protocolo SIP (Session Initiation Protocol) [11] (que es el protocolo de señalización utilizado en IMS), la pregunta inmediata es qué tamaño medio tendrán estos paquetes que deben ser procesados por una aplicación y cuál será su tasa de generación. Considerando, en el caso peor, tamaños medios de mensajes de SIP pequeños del orden de unos *465 bytes* para el caso de mensajes INVITE o *388 bytes* para mensajes OK, nos situamos en estas pruebas siempre con tasas no superiores a 40 Mbps. En situaciones más complejas, donde las cabeceras de los mensajes SIP pueden ser más extensas, se debe de considerar, como se recoge en [12], que los mensajes de SIP pueden estar entre los *838* y *1024 bytes*. Lo que se indica como tal en [11], que estandariza SIP, es que la longitud de sus mensajes no superen el valor de la MTU (Maximum Transmission Unit) si es previamente conocida, o 1200 bytes en su defecto.

De todas formas, una prueba interesante es cambiar el equipo que realiza las funciones del RGW por otro más potente de *sobremesa*, para comprobar el efecto del procesador en estas pruebas. Para ello se utilizó un ordenador Pentium 4 con 2,4 GHz y 512 Mbytes de memoria RAM (igual que el ordenador compacto) en vez del *compacto* que se utilizará como RGW. En la Fig. 8 se pueden observar los resultados obtenidos. En esta gráfica también se han eliminado los resultados del escenario que se utiliza IPTables como NAT.

Cabe destacar la mejora con respecto a las pruebas realizadas con el ordenador *compacto* por lo que se puede concluir que el procesador es un elemento que definitivamente influye en el caso de relegar el procesamiento de las tramas al nivel de aplicación.

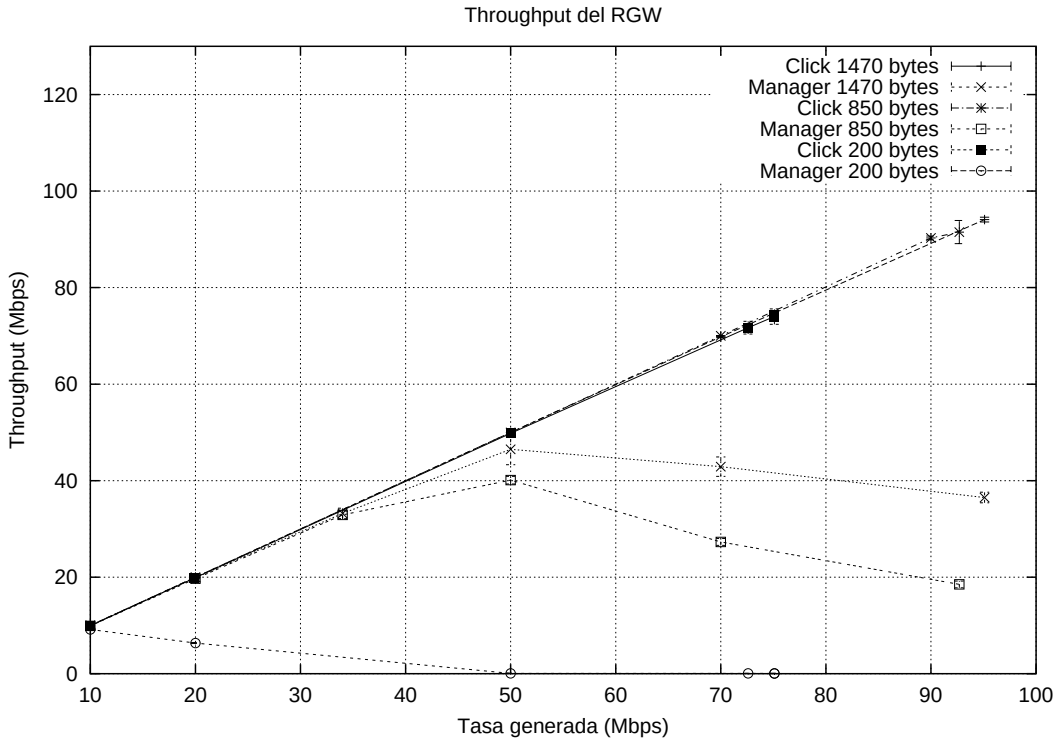


Figura 7: Resultados usando el ordenador compacto.

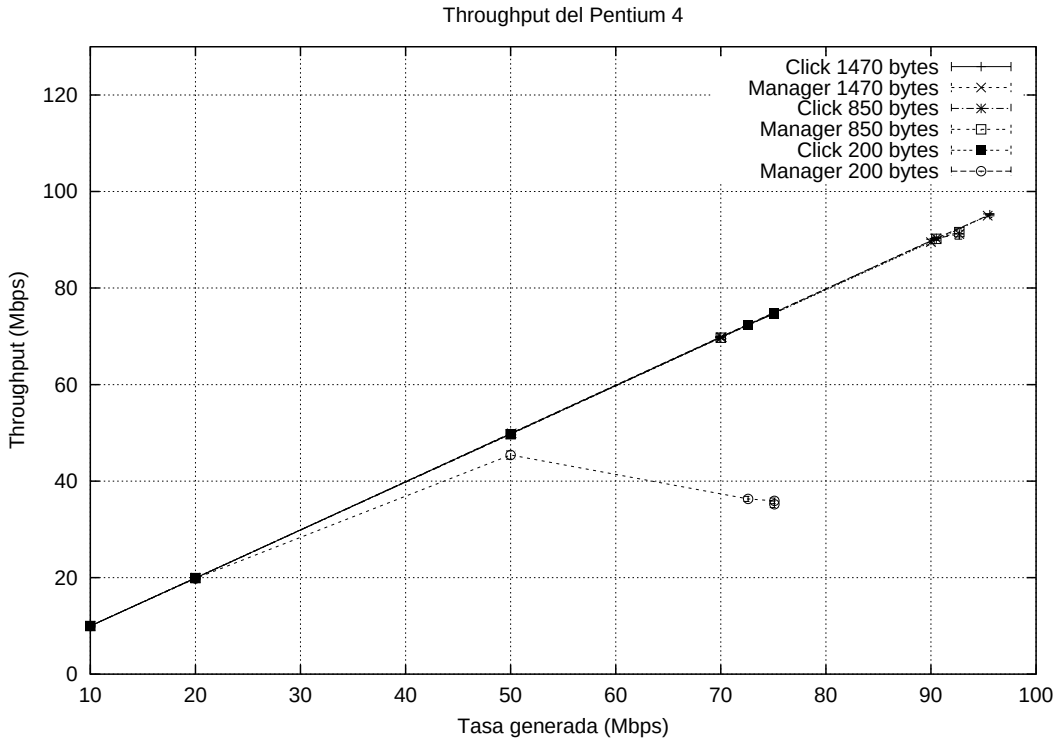


Figura 8: Resultados usando el ordenador de sobremesa.

5. Conclusiones

En el artículo se han presentado las diferentes pruebas realizadas con la plataforma de desarrollo de routers modulares Click con el objetivo de validar su utilización dentro del proyecto MUSE.

Este artículo propone la utilización de un modelo híbrido que se ha probado en diferentes escenarios en donde se han planteado escenarios reales utilizando los equipos hardware que se planean utilizar en el prototipo final.

Entre los diferentes resultados cabe destacar el

obtenido al probar la carga soportada en las diferentes configuraciones en donde queda patente la importancia que tiene un procesador potente en el equipo RGW y el tamaño de los paquetes que deben ser procesados en el nivel de aplicación, es decir, de los paquetes que el nivel de Click enviará hacia el nivel de aplicación para ser tratados por el Manager. En el artículo se ha demostrado como incluso con un ordenador que no tenga demasiados recursos, será posible tratar los mensajes de señalización que habitualmente son mensajes pequeños soportando un caudal más que suficiente. Una vez terminado el proceso de señalización, el tráfico de usuario será tratado ya a nivel de Click y en las diferentes pruebas se ha comprobado que a este nivel no hay perjuicio apreciable de prestaciones.

A la hora de tratar los mensajes al nivel de aplicación, otro resultado importante es el obtenido al comparar los elementos de Click *ToHost* y *ToHostSniffers* (mecanismos habituales para transferir los datos a niveles superiores). Las librerías *jpcap* de Java estudiadas no presentan ninguna ventaja con respecto a los *sockets* normales de Java, por lo que se ha estimado que el mecanismo habitual de envío de estos datos (*toHost*) es suficiente como para tratar estos flujos y además presenta un menor nivel de complejidad. Queda como trabajo futuro la realización de las mismas pruebas, esta vez con un Manager programado en lenguaje C, donde se pueden utilizar las funciones de la librería *libpcap*.

Agradecimientos

Este artículo ha sido financiado parcialmente por la Comisión Europea a través del proyecto MUSE.

Referencias

[1] Multi service access everywhere (muse) european project. [Online]. Available: http://www.ist-muse.org/

[2] E. Kohler, R. Morris, B. Chen, J. Jannotti, and M. Kaashoek, "The click modular router.acm transactions on computer systems," *ACM Transactions on Computer Systems*, vol. 18, no. 3, pp. 263–297, 2000.

[3] *IP Multimedia Subsystem (IMS)*, 3GPP TS 23.228 V6.6.0, Rev. Stage 2 (Release 6), 2004.

[4] "Release 1: Release definition," TISPAN: Draft ETSI, 2004.

[5] Lex system. [Online]. Available: http://www.lex.com.tw:8080/ home.htm

[6] Página web de click. [Online]. Available: http://www.pdos.lcs.mit.edu/click/

[7] N. R. G. Lawrence Berkeley National Labs. libpcap. [Online]. Available: http://www.tcpdump.org/

[8] Jpcap: Java package for packet capture. [Online]. Available: http://netresearch.ics.uci.edu/kfujii/jpcap/doc/

[9] Jpcap: a network packet capture library. [Online]. Available: http://jpcap.sourceforge.net/

[10] Iperf. [Online]. Available: http://dast.nlanr.net/Projects/Iperf/

[11] J. Rosemberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: Session Initiation Protocol," IETF RFC 3261, 2002.

[12] H. Schulzrinne and J. Rosemberg, "The session initiation protocol: Internet-centric signalling," *IEEE Communications Magazine*, vol. 38, no. 10, 2000.

Una aplicación del RFC 4038: la metodología de transición a IPv6 MENINA

Eva M. Castro
Dpto. de Automática - UAH
Ctra Madrid-Barcelona, Km 33,6
28871 - Alcalá de Henares
eva@aut.uah.es

Pedro de las Heras
Dpto de Inf. Est. y Telemática - URJC
C/ Tulipán s/n
28933 - Móstoles
pheras@gsyc.escet.urjc.es

Tomás de Miguel, Santiago Pavón
Dpto. de Ing. y Sist. Telemáticos - UPM
Ciudad Universitaria s/n
28040 - Madrid
{tmiguel,santiago}@dit.upm.es

Abstract *Recently, a new RFC has been published: "RFC 4038: Application Aspects of IPv6 Transition". This RFC, co-authored by one of the authors of this paper, explains the close connection between the applications and networks in the IPv6 transition process. Since the publication of the IPv6 specification and its subsequent multiple revisions, people have been worried about the transition of network devices. The technology is ready for upgrading to IPv6 and a lot of transition mechanisms have been defined in the last years. However, the transition of applications has been left behind. This paper provides some rationale about the IPv6 application transition recommendations of the RFC 4038, and later describes the MENINA (MEthodology for the New IPv6 Network Adaptation) methodology. This methodology helps and guides users in the transition process, focusing not only on the network infrastructure, but also on the applications.*

1. Introducción

Al principio de la década de los 90, las previsiones de crecimiento de la red Internet y la aparición de necesidades relacionadas con los nuevos servicios sirvieron de detonante para la revisión del protocolo IPv4.

La nueva versión IPv6 aborda el problema del tamaño de la red proponiendo una nueva estructura de direcciones y un nuevo esquema de encaminamiento. Además, el nuevo modelo ofrece capacidades para resolver las necesidades de la nueva generación de servicios, integrando soluciones de seguridad, movilidad y calidad de servicio.

Debido a que todos los nodos de la red no pueden estar preparados para IPv6 al mismo tiempo, no se puede hacer una sustitución de un protocolo por otro de forma instantánea, por lo que el proceso de transición está siendo lento lo que hará que durante un largo periodo de tiempo vayan a tener que convivir ambos protocolos.

La estrategia para adoptar IPv6 como protocolo de red dentro de una organización es una decisión que tiene repercusiones en todo el modelo de comunicaciones de una organización. Un cambio de esta magnitud requiere la revisión de todos los sistemas relacionados con el intercambio de información a través de la red.

Los grupos de trabajo de IPv6 dentro del IETF, IPv6 Operations IETF Working Group (v6ops) y antiguamente Next Generation Transition IETF Working Group (ngtrans), han propuesto una serie de mecanismos de transición para permitir el correcto funcionamiento de redes heterogéneas, IPv4 e IPv6. La mayor parte del trabajo relacionado con la coexistencia de IPv4 e IPv6 se ha realizado en este campo.

Se han realizado algunos estudios [2] [3] [4] en el campo de la transición para esclarecer el funcionamiento de la gran cantidad de mecanismos de transición existentes. Normalmente, estos trabajos se limitan a resaltar las características más importantes de cada uno de los mecanismos focalizando el problema en la red de distribución.

Desde el IETF se apoyan estudios relacionados con los casos típicos de transición de redes según el sector donde estén desplegadas: redes de ISP [5], redes no gestionadas [6], redes empresariales [7] y redes 3rd Third Generation Partnership Project (3GPP) [8].

Todos estos estudios anteriores se centran en el cambio de versión de la pila IP, es decir, las funciones de direccionamiento y encaminamiento propias del nivel de red. Sin embargo, la incompatibilidad de las dos versiones del protocolo IP alcanza a las aplicaciones.

Las aplicaciones diseñadas para IPv4 no pueden operar directamente con el protocolo IPv6 y por tanto deben ser revisadas y adecuadamente adaptadas. La interfaz que utilizan las aplicaciones, API (Application Program Interface), para la

programación de funciones de intercambio de información con otras entidades, hace visible a las aplicaciones los detalles de la versión de IP utilizada, provocando que éstas sean dependientes de la versión de IP.

Algunos estudios [9] [10] previos al RFC 4038 analizan el problema de la transición de las aplicaciones, considerando las dependencias que existen con la versión de IP dentro del código fuente. Sin embargo, la publicación del RFC 4038 por parte del IETF define unas pautas de transición desde el punto de vista de las aplicaciones e incluye código fuente de ejemplo, que servirá como punto de referencia para los desarrolladores.

Este artículo está divido en dos partes. En la primera parte se presentan las principales aportaciones realizadas por el RFC 4038 [11], al que hemos contribuido los autores desde sus orígenes como *draft* del IETF, y que nos han permitido conocer la necesidad de plantear la transición como un proceso en el que intervienen tanto las redes como las aplicaciones. En la segunda parte del artículo se describe la metodología MENINA, la cuál plantea soluciones de coexistencia entre IPv4 e IPv6, tanto desde el punto de vista de las redes como desde las aplicaciones, siguiendo las recomendaciones del RFC 4038.

2. RFC 4038

En el RFC 4038 se describe la mejor estrategia de transición para las aplicaciones basándose en la descripción de un conjunto de escenarios típicos de transición. En estos escenarios se contemplan básicamente dos posibilidades desde el punto de vista de las aplicaciones. Por un lado, la instalación de mecanismos de transición que permitan la interoperabilidad de código fuente IPv4 con las nuevas redes IPv6. Y por otro, la modificación de las aplicaciones para que funcionen con IPv6 de forma nativa. Para ello, se describen detalladamente todas las posibles dependencias de las aplicaciones con la versión de IP.

2.1. Escenarios de transición

En numerosas ocasiones se puede conseguir la coexistencia a través de nodos con doble pila o duales, es decir, nodos que tienen instalada tanto la pila IPv4 como la pila IPv6, y que utilizarán uno u otro protocolo dependiendo de las necesidades requeridas. A continuación se describen los escenarios típicos de transición en los que tendremos que proporcionar coexistencia entre aplicaciones y redes heterogéneas.

- Aplicaciones IPv4 en nodos duales

En este escenario, aunque los nodos están preparados para funcionar con ambos protocolos, las aplicaciones sólo pueden funcionar con IPv4. En este escenario el problema de interoperabilidad surge si la aplicación IPv4 necesita comunicarse utilizando IPv6. Para permitir a estas aplicaciones intercambiar tráfico IPv6, sería necesario utilizar alguno de los mecanismos de transición de nodo propuestos por el IETF: BIS [12] y BIA [13]. Aunque siempre es recomendable tratar de modificar las aplicaciones IPv4 para que funcionen de forma nativa con IPv6, estos mecanismos pueden servir de ayuda en casos especiales en los que no se disponga del código fuente de las aplicaciones o de la licencia apropiada para su modificación.

Estos mecanismos se instalan en el nodo que ejecuta la aplicación IPv4 y proporcionan direcciones IPv4 ficticias a la aplicación para que pueda establecer sus comunicaciones como si la entidad remota fuera IPv4 y ellos se encargan de realizar la traducción entre ambos protocolos dentro del nodo. La diferencia fundamental entre BIS y BIA radica en el nivel en el que se realiza la traducción entre IPv4 e IPv6. En el caso de BIA se realiza en el API (Application Program Interface) que utilizan las aplicaciones y en el caso de BIS se lleva a cabo entre la implementación IP y el gestor de la tarjeta de red.

- Aplicaciones IPv6 en nodos duales

Las aplicaciones IPv6 son aplicaciones que trabajan con IPv6 de forma nativa. El problema de interoperabilidad surge si estas aplicaciones necesitan comunicarse utilizando IPv4.

La mayoría de los nodos duales permiten un modo de funcionamiento en el que las aplicaciones sólo IPv6 pueden intercambiar paquetes IPv4. Este modo de funcionamiento se consigue cuando las aplicaciones IPv6 utilizan direcciones IPv6 especiales, las *direcciones IPv4-mapped* que tienen el siguiente formato: *"::FFFF:x.y.z.w"*, donde *"x.y.z.w"* es una dirección IPv4 válida de un nodo remoto. Al utilizar este tipo de direcciones, la implementación de la doble pila del nodo realiza la traducción adecuada para que el protocolo utilizado con la aplicación sea IPv6 y el protocolo utilizado con la entidad remota sea IPv4.

Desafortunadamente, no todas las implementaciones de la doble pila en los sistemas operativos actuales permiten este modo de funcionamiento.

- Aplicaciones duales en nodos duales

Al modificar el código fuente de una aplicación para que funcione con IPv6, lo más recomendable es diseñar estas modificaciones para que la aplicación pueda funcionar con IPv4 y con IPv6, lo que hemos denominado aplicación dual. Este escenario en el que las aplicaciones duales se ejecutan en nodos duales es el que permite una mayor interoperabilidad, ya que tanto las aplicaciones como el propio nodo pueden funcionar con ambas versiones del protocolo IP. En estos casos, las aplicaciones deberían contemplar como comportamiento por defecto utilizar comunicaciones IPv6 y si éstas fallan utilizar la versión IPv4.

- Aplicaciones duales en nodos IPv4

Una vez que las aplicaciones IPv4 se han modificado para que sean duales, puede llegar a ocurrir que tengan que continuar ejecutándose en nodos en los que sólo se tenga instalada la pila IPv4. Este escenario puede presentarse en aquellos casos en los que no se quieran mantener 2 versiones de la misma aplicación, una versión IPv4 para los nodos sólo IPv4 y otra dual para el resto de los nodos. Para evitar posibles problemas de interoperabilidad es necesario que en el desarrollo de aplicaciones duales no se realice ninguna suposición sobre la versión del protocolo IP que van a utilizar para comunicarse.

2.2. Dependencias de las aplicaciones con la versión de IP

La interfaz que utilizan las aplicaciones, API , para la programación de funciones de intercambio de información con otras entidades, hace visible a las aplicaciones los detalles de la versión de IP utilizada, provocando que éstas sean dependientes de la versión de IP. A continuación se describen las dependencias más habituales con la versión de IP que pueden encontrarse en el código fuente de las aplicaciones.

- Formato de presentación de las direcciones Muchas aplicaciones utilizan direcciones IP para identificar los nodos remotos con los que quieren establecer una conexión. Estas direcciones normalmente se proporcionan a las aplicaciones utilizando el formato de presentación, es decir, como una cadena de caracteres, que la aplicación tendrá que analizar para convertirla en la estructura de datos adecuada.

Este análisis depende de la versión de IP que estemos utilizando. Las direcciones IPv4 están formadas por 4 números decimales separados por el carácter "." y las direcciones IPv6 están formadas por 8 palabras de 16 bits en formato hexadecimal separadas por el carácter ":".

Además, normalmente las direcciones IPv4 en formato de presentación son más cortas que las direcciones IPv6. Por tanto, la cantidad de memoria reservada por las aplicaciones IPv4 para almacenar una dirección IPv4 en formato de presentación no es suficiente para almacenar una dirección IPv6.

Por tanto, es necesario revisar el formato de presentación de las direcciones IP para poder soportar ambas versiones del protocolo.

- API del nivel de transporte

Las aplicaciones hacen uso de las funciones de programación que proporciona la interfaz de nivel de transporte para establecer comunicaciones. En el caso más general, el hecho de adaptar una aplicación a IPv6 requiere un examen exhaustivo del código fuente de las aplicaciones buscando los siguientes aspectos relacionados con el API de comunicaciones:

- Almacenamiento de la información de red, en particular, las direcciones IP. Las direcciones IPv6 ocupan 128 bits frente a los 32 bits de las direcciones IPv4.

- Funciones de conversión de las direcciones. Estas funciones se encargan de convertir las estructuras de datos que almacenan las direcciones IP en el formato presentación de las mismas.

- Funciones para la apertura, cierre y gestión de las comunicaciones. Estas funciones se encargan de gestionar todo lo relativo a las comunicaciones para el intercambio de información entre las aplicaciones.

- Opciones de configuración de red. Estas opciones especifican diferentes modos de comunicación y normalmente dependen de la versión de IP utilizada.

- Resolución de nombres y direcciones

La tarea de resolución entre nombres y direcciones IP es independiente de las aplicaciones. Éstas utilizan un conjunto de funciones que permiten consultar la asociación existente entre direcciones IP y nombres de máquina.

Las funciones para la consulta de la resolución de nombres y direcciones IPv4 no funcionan con IPv6 y será necesario utilizar las nuevas funciones que se hayan desarrollado en el API para la nueva versión del protocolo IP.

- Otras dependencias específicas

Aunque normalmente las dependencias respecto de la versión IP son las citadas previamente, existen algunas aplicaciones que muestran otro tipo de dependencias debidas a la utilización de las direcciones IP con significados específicos para dichas aplicaciones. Las más habituales son:

- Intercambio de las direcciones IP como datos del nivel de aplicación. Por ejemplo, el protocolo FTP.

- Utilización de las direcciones IP en el nivel de aplicación para identificar nodos, usuarios u otras características de la aplicación. En algunos casos incluso se registran y se almacenan para utilizarlas dentro de sistemas de autenticación.

- Utilización de las direcciones IP como parámetros de entrada para la aplicación. En estos casos, el intérprete de los parámetros de entrada dependerá de la versión de IP utilizada.

3. MENINA

Siguiendo las recomendaciones descritas en los escenarios típicos de transición propuestos en el RFC 4038 y el estudio de todos los mecanismos de transición de red especificados hasta ahora, se ha definido la metodología MENINA [14] para permitir guiar al usuario en el proceso de transición a IPv6, partiendo del escenario de red concreto donde desea realizar el despliegue de la nueva versión del protocolo IP.

El problema que aborda la metodología MENINA es la transición gradual desde el punto de vista global del funcionamiento de las aplicaciones y las redes en escenarios heterogéneos. Denominamos escenarios heterogéneos a aquellos escenarios donde se mezclan aplicaciones heterogéneas, IPv4 e IPv6, y redes también heterogéneas, IPv4 e IPv6. Hasta ahora, este problema había sido abordado de manera independiente, tratando de resolver por una parte los problemas de coexistencia de diferentes redes y por otra, los problemas de interoperabilidad entre aplicaciones.

La metodología MENINA parte de un conjunto de datos iniciales que definen el problema concreto de un usuario, es decir, una descripción del tipo de aplicaciones que se desea que funcionen sobre un escenario de red determinado. La metodología ofrece como resultado el conjunto de soluciones de transición que garantizan la comunicación entre las aplicaciones descritas inicialmente. Las soluciones de transición están basadas en la selección de los mecanismos de transición adecuados para el escenario inicial de red y en las consideraciones para obtener aplicaciones compatibles con ambos protocolos.

Además, la metodología incluye un modelo de especificación de los costes asociados a cada solución que permite obtener una ordenación de las soluciones siguiendo un criterio de costes. El usuario de la metodología deberá proporcionar una estimación de los mismos en su caso concreto de transición. Esta estimación no se puede automatizar ya que es muy dependiente de la situación particular del usuario, quién tendrá que valorar los requisitos necesarios para el despliegue de cada una de las soluciones.

Por tanto, la metodología MENINA proporciona un conjunto de soluciones de transición ordenadas según el criterio de costes de despliegue asociados a cada una de ellas.

Para la especificación formal de MENINA se ha utilizado una notación genérica que permite describir formalmente cada uno de los componentes que participan en el proceso de transición y las reglas de transformación necesarias que proporcionarán las soluciones de transición. Esta representación formal ha permitido el desarrollo de una herramienta asociada a MENINA que aplica la metodología de forma automática.

A continuación se describen los componentes básicos que definen la metodología MENINA: el modelo conceptual, las etapas de la metodología de transición y la herramienta.

3.1. Modelo conceptual

Primeramente es necesario identificar aquella parte del mundo real que es relevante para el proceso de la transición, es decir, los componentes de un escenario global de red que se ven afectados por el cambio de la versión del protocolo IP. Todos estos componentes, junto con las reglas y convenciones que se usarán para describir su comportamiento dentro de un escenario de transición, formarán el modelo conceptual de la metodología.

Dado que la mayoría de las aplicaciones utilizan comunicación unicast, el modelo conceptual considera los siguientes elementos: dos aplicaciones que intercambian información a través de una infraestructura de red siguiendo el modelo de comunicación uno-a-uno. Por tanto, el modelo conceptual describe los siguientes componentes (véase la figura 1) junto con su representación formal que será utilizada en la metodología MENINA:

- Servicios

Una aplicación es un programa que se ejecuta dentro de un nodo y que realiza una función específica directamente para el usuario, proporcionándole un servicio. La forma más habitual de proporcionar un servicio al usuario es la utilización del modelo de comunicaciones cliente/servidor. Por una parte, existe una aplicación funcionando como servidor, ofreciendo un servicio. Y, por otra, existe una aplicación funcionando como cliente, solicitando dicho servicio. Ambas aplicaciones pueden ser independientes, sin embargo, comparten un protocolo de comunicación que las permite intercambiar información y proporcionar el servicio al usuario final.

No todas las aplicaciones que se ejecutan en los nodos son importantes para el proceso de transición. Las aplicaciones relevantes son aquéllas que utilizan la red como medio de comunicación con otras aplicaciones.

Desde el punto de vista del proceso de transición, las aplicaciones se clasifican según la versión de IP para la que fueron desarrolladas: aplicaciones sólo-IPv4 (A_4), aplicaciones sólo-IPv6 (A_6) y aplicaciones duales (A_d).

Dentro de la metodología MENINA representaremos todos los tipos de aplicaciones utilizando los elementos del conjunto $\mathcal{A} = \{A_4, A_6, A_d\}$.

- Nodos

Un nodo es un dispositivo que permite la ejecución de aplicaciones y servicios de usuario. En el proceso de transición los elementos relevantes de un nodo [15] son aquéllos que realizan funciones concretas relacionadas con la transmisión de información.

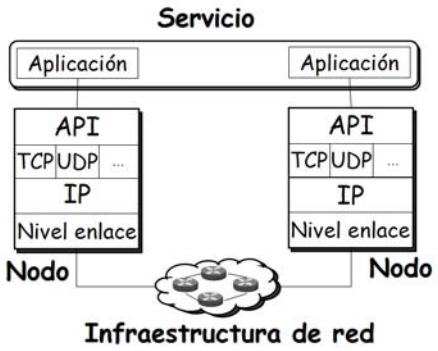


Figura 1: Modelo Conceptual.

Desde el punto de vista del proceso de transición, el aspecto clave de un nodo es su capacidad para intercambiar información entre las aplicaciones locales y la red IP. En concreto, para la metodología son interesantes los siguientes aspectos:

- La pila IP: nodo sólo-IPv4, nodo sólo-IPv6 y nodo dual.
- Los mecanismos de transición que proporcionan interoperabilidad al nodo: la capacidad de utilizar las direcciones especiales *IPv4-mapped* y los mecanismos BIS y BIA.

Dentro de la metodología MENINA representaremos todos los tipos de nodo utilizando el conjunto $\mathcal{N}$.

- Infraestructura de red

Definimos la infraestructura de red como la lista de nodos intermedios que permiten comunicarse entre sí a dos nodos terminales.

Desde el punto de vista de la transición, solamente estamos interesados en los nodos intermedios que usan el protocolo de red IP. Por lo tanto sólo se tienen en cuenta los nodos encaminadores intermedios. Los aspectos que nos interesa modelar de los nodos intermedios son los siguientes:

- La versión IP que utilizan: sólo-IPv4, sólo IPv6 y encaminadores duales.
- Algunas funciones de los encaminadores que influyen en las soluciones de transición: IPv4 NAT y protocolo de traducción IPv4/IPv6.

Representamos todos los tipos de nodos intermedios o encaminadores utilizando el conjunto $\mathcal{R}$. Utilizamos el operador $\leftrightarrow$ para representar la conexión entre nodos intermedios:

$$r_i \leftrightarrow r_j, , \quad r_i, r_j \in \mathcal{R}$$

La infraestructura de red es una lista de n nodos intermedios que pertenecen a $\mathcal{R}^n$, perteneciendo cada encaminador a $\mathcal{R}$.

3.2. Etapas de la metodología

Se define un escenario de transición, S, como un par de aplicaciones, $a1$ y $a2$, que corren en nodos terminales $n1$ y $n2$, que están conectados por una lista de encaminadores, rl:

$$\begin{aligned} S =& < a1, n1, rl, n2, a2 > \\ & S : transition\ \ scenario \\ & a1, a2 \in \mathcal{A} \\ & n1, n2 \in \mathcal{N} \\ & rl \in \mathcal{R}^n, , \quad rl = r_i(\leftrightarrow r_j)* \end{aligned}$$

La lista de encaminadores rl se compone de uno o más nodos intermedios conectados a través del operador de conexión $\leftrightarrow$. La infraestructura de red IP degenerada conecta los nodos terminales directamente sin ningún nodo intermedio, $rl = \{\}$.

Por lo tanto, S describe los elementos que participan en el proceso de comunicación entre aplicaciones desde el punto de vista de la transición a IPv6. Durante el periodo de transición, las aplicaciones heterogéneas, los nodos terminales y los encaminadores deben coexistir simultáneamente. Debido a que los dos protocolos son incompatibles, deberemos estudiar la posibilidad de comunicaciones en el escenario S.

La metodología MENINA, partiendo de un conjunto de escenarios de transición inicial, un conjunto de S, analiza su disponibilidad de comunicaciones, y devuelve transformaciones de los escenarios que resuelven los problemas de comunicaciones. Estas transformaciones se basan en la instalación de mecanismos de transición y en la adaptación de aplicaciones.

La metodología MENINA guía el proceso de transición definiendo las siguientes etapas: normalización del escenario de red, búsqueda de soluciones, y evaluación de soluciones. A continuación analizamos cada una de las etapas.

3.2.1. Normalización del escenario de red

Esta etapa parte de un escenario inicial, S, y transforma la infraestructura de la red inicial, rl, en una red normalizada. La normalización proporciona una representación algebraica del escenario inicial a través de *zonas de encaminamiento y operadores de conexión*.

Las *zonas de encaminamiento homogéneas* permiten enlazar los encaminadores de acuerdo a su comportamiento respecto al encaminamiento IP: zonas sólo-IPv4, sólo-IPv6 y duales. Representamos todos los tipos de zonas utilizando el conjunto $\mathcal{Z}$.

Las zonas se conectan mediante *operadores de conexión* que se caracterizan por sus capacidades: encaminamiento IP, NAT IPv4 y traducción de protocolos IPv4/IPv6.

3.2.2. Búsqueda de soluciones

La etapa de búsqueda de soluciones parte del escenario normalizado producido por la etapa 1 y aplica todas las transformaciones posibles para obtener escenarios transformados en los que la comunicación puede producirse. Las transformaciones se basan en la instalación de mecanismos de transición y en la adaptación de aplicaciones.

Para implementar esta etapa la metodología necesita que se proporcione una representación de los mecanismos de transición y la adaptación de aplicaciones utilizando la anterior notación. A continuación se muestran los tipos de transformaciones:

- Transformaciones entre zonas de encaminamiento, basadas en la configuración de túneles entre zonas: túneles manuales [16], broker de túneles [17] y 6to4 [18].
- Transformaciones entre una zona y un nodo terminal, basadas en la configuración de túneles entre una zona y un nodo terminal: túneles manuales, broker de túneles, DSTM [19], Teredo [20] e ISATAP [21].
- Las transformaciones entre los dos nodos terminales, basado en la configuración de túneles entre los nodos terminales: túneles manuales e ISATAP.
- Transformaciones de los nodos intermedios, basadas en la instalación de traducción de protocolos IPv4/IPv6: SIIT, NATP-PT y TRT.
- Transformaciones de los nodos terminales, basadas en la instalación de mecanismos en los nodos terminales: direcciones IPv6 mapeadas en direcciones IPv4, BIS [12] y BIA [13].
- Transformaciones de las aplicaciones [11], basadas en las modificaciones del código fuente que se necesitan para proporcionar aplicaciones duales.

Se han definido estos mecanismos de transición como reglas de transformación utilizando la representación normalizada. Al ser esta representación genérica, se pueden añadir nuevos mecanismos de transición que aparezcan en el futuro representando su comportamiento como nuevas reglas de transformación.

3.2.3. Evaluación de soluciones

La salida de la etapa 2 es el resultado de analizar todas las transformaciones que hacen posible la comunicación. Esta etapa ordena las soluciones atendiendo a criterios económicos.

La estimación de costes de las transformaciones sólo puede llevarla a cabo la organización que está realizando el proceso de transición. La naturaleza de las transformaciones depende mucho de la situación específica de la organización: los contratos con proveedores de hardware, licencias de las aplicaciones, personal cualificado disponible para llevar a cabo las tareas, etc.

En esta etapa se le presenta a la organización una lista detallada de los cambios requeridos por cada transformación para que pueda realizar la estimación de costes. La etapa de evaluación de soluciones utiliza esta información y proporciona la solución más barata que mejor se adecúa al conjunto inicial de escenarios.

3.3. Herramienta

La representación algebraica de los escenarios y las transformaciones de transición se ha realizado utilizando el lenguaje de descripción formal LOTOS [22], lo que facilitó la definición de la herramienta MENINA. Esta herramienta aplica automáticamente la metodología MENINA. La figura 2 muestra la arquitectura de la herramienta.

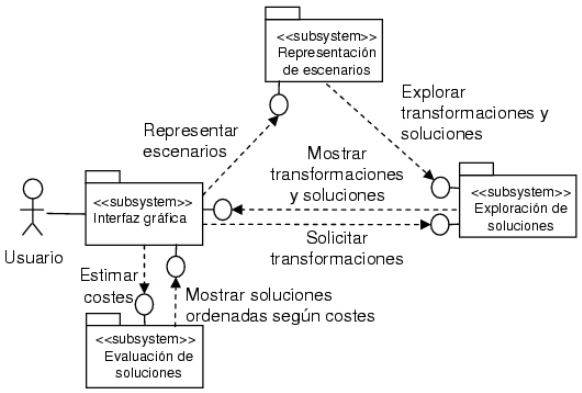


Figura 2: Arquitectura de la herramienta MENINA.

Además de la interfaz gráfica, se han definido tres subsistemas básicos en la arquitectura, correspondiendo con las etapas de la metodología MENINA:

- Interfaz gráfica. Este subsistema permite realizar la captura de los escenarios iniciales y mostrar la representación de las soluciones al usuario.
- Representación de escenarios. Este subsistema realiza la etapa 1 de MENINA; se encarga de normalizar el escenario de red inicial utilizando la representación algebraica.
- Exploración de soluciones. Este subsistema aplica todas las transformaciones definidas en la etapa 2 de MENINA, basándose en los mecanismos de transmisión y de adaptación de aplicaciones. El resultado es un conjunto de escenarios en los que la comunicación puede tener lugar.

- Evaluación de soluciones. Este subsistema ordena las soluciones generadas por la etapa, de acuerdo a los criterios económicos que proporciona el usuario.

4. Aplicación de MENINA

En esta sección se presenta un caso de estudio al que se ha aplicado la metodología MENINA. Este caso es el de una red no gestionada, que pertenece a una oficina pequeña o a una red doméstica. La configuración más habitual para una red IPv4 de este tipo está formada por un router IPv4 conectado al ISP IPv4 y un conjunto de nodos IPv4 conectados dentro de la misma subred. Véase la figura 3.

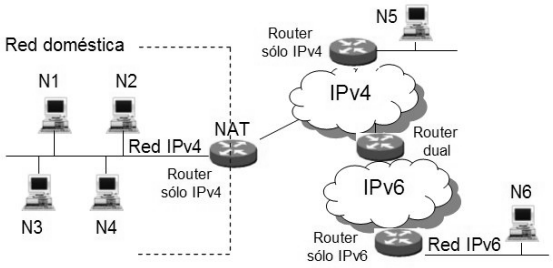


Figura 3: Caso de estudio: red no gestionada.

El propietario de la red marcará como objetivo el funcionamiento de diferentes patrones de comunicación entre las aplicaciones. Supongamos, por ejemplo, que se marca como objetivo el funcionamiento de aplicaciones IPv6 en los nodos $N1$ y $N4$ con aplicaciones IPv6 externas en una red IPv6, en el nodo $N6$. Este escenario se representa a través del siguiente patrón de comunicaciones:

$$S =< A_6, N_4, rl, N_6, A_6 >$$

$$donde\ \ rl\ =\ R_4^t(\leftrightarrow R_4)* \leftrightarrow R_d(\leftrightarrow R_6)* \leftrightarrow R_6$$

Etapa 1: Normalización del escenario de red

La normalización transforma el escenario inicial en una representación basada en la conexión de zonas a través de operadores de conexión:

$$A_6 \diamond N_4 \leftrightarrow Z_4 \odot_4 Z_4 \oplus_d Z_6 \leftrightarrow N_6 \diamond A_6$$

donde todos los R_4 se agrupan en una Z_4, todos los R_6 se agrupan en una Z_6, $\odot_4$ es un operador que conecta dos zonas a través de IPv4 con capacidad de realizar NAT y $\oplus_d$ es un operador que conecta las dos zonas con ambos protocolos.

Etapa 2: Búsqueda de soluciones

MENINA aplica sistemáticamente las reglas de transformación que representan a los mecanismos de transición y a la adaptación de aplicaciones, encontrando las siguientes soluciones:

- Solución 1: túneles para atravesar zonas $Z_4 \odot_4 Z_4$. Estas soluciones pueden implementarse utilizando alguno de los siguientes mecanismos: túnel IPv6 dentro de UDP IPv4, túnel configurado manualmente, broker de túneles o Teredo. Este tipo de soluciones requieren convertir el nodo N_4 en la zona IPv4 a N_d.

$$A_6 \diamond \overbrace{N_4 \leftrightarrow Z_4 \odot_4 Z_4 \oplus_d Z_6}^{N_d \leftrightarrow Z_6} \leftrightarrow N_6 \diamond A_6 \ \equiv$$

$$A_6 \diamond N_d \leftrightarrow Z_6 \leftrightarrow N_6 \diamond A_6$$

Todas estas soluciones desde el punto de vista de la transición, convierten la conexión de esas zonas heterogéneas en una única Z_6.

- Solución 2: direcciones *IPv4-mapped IPv6* y traducción de protocolos. El uso de estas direcciones requiere convertir el nodo N_4 en N_d, así la A_6 podrá intercambiar tráfico IPv4 en Z_4. La traducción de protocolos se representa por el operador $\boxplus_d$.

$$A_6 \diamond \overbrace{N_4}^{N_{d,MAP}} \leftrightarrow Z_4 \odot_4 Z_4 \overbrace{\oplus_d}^{TP} Z_6 \leftrightarrow N_6 \diamond A_6 \ \equiv$$

$$A_6 \diamond N_{d,MAP} \leftrightarrow Z_4 \odot_4 Z_4 \ \boxplus_d \ Z_6 \leftrightarrow N_6 \diamond A_6, ,$$

- Solución 3: adaptación de una aplicación y traducción de protocolos. Adaptamos la aplicación A_6 para que sea dual y pueda intercambiar tráfico IPv4 en Z_4 y utilizamos traducción de protocolos para convertir el tráfico a IPv6.

$$\overbrace{A_6}^{A_d} \diamond N_4 \leftrightarrow Z_4 \odot_4 Z_4 \overbrace{\oplus_d}^{TP} Z_6 \leftrightarrow N_6 \diamond A_6 \ \equiv$$

$$A_d \diamond N_4 \leftrightarrow Z_4 \odot_4 Z_4 \boxplus_d Z_6 \leftrightarrow N_6 \diamond A_6$$

- Solución 4: adaptación de una aplicación y utilización de un túnel IPv4 dentro de IPv6 para atravesar la Z_6. Adaptamos la aplicación a dual para que pueda intercambiar tráfico IPv4 dentro de Z_4. Para la configuración de este túnel es necesario convertir N_6 en N_d dentro de Z_6 y puede realizarse manualmente, con broker de túneles o con DSTM.

$$\overbrace{A_6}^{A_d} \diamond N_4 \leftrightarrow Z_4 \odot_4 \overbrace{Z_4 \oplus_d Z_6 \leftrightarrow N_6}^{Z_4 \leftrightarrow N_d} \diamond A_6 \ \equiv$$

$$A_d \diamond N_4 \leftrightarrow Z_4 \odot_4 Z_4 \leftrightarrow N_d \diamond A_6$$

Etapa 3: Evaluación de soluciones

Dentro de la etapa de evaluación de soluciones, el responsable del escenario de red tiene que realizar la estimación de costes en función de las soluciones aportadas por la segunda etapa. En el caso concreto que estamos analizando, el usuario de una red doméstica no tiene capacidad para realizar cambios de configuración fuera del ámbito de su red, por tanto, sólo nos queda la solución 1 con Teredo o el broker de túneles para atravesar $Z_4 \odot_4 Z_4$. Sin detallar el proceso de estimación

de costes y evaluación de soluciones, para este caso podemos suponer que el usuario de la red doméstica podría utilizar alguno de los brokers de túneles que se encuentran disponibles de forma gratuita y por tanto, preferir esta solución frente a Teredo.

5. Conclusiones

Tal como refleja el recientemente publicado RFC 4038 en el que hemos participado los autores, es necesario tener en cuenta en la transición de IPv4 a IPv6 la importancia tanto de las redes como de las aplicaciones.

La metodología MENINA se ha diseñado para contribuir a la resolución del proceso de transmisión IPv6. MENINA proporciona soluciones a escenarios de red específicos en los que las aplicaciones heterogéneas que corren en nodos terminales heterogéneos están intercambiando información a través de una infraestructura de red. MENINA explora sistemáticamente el espacio de transformaciones. El RFC 4038, la metodología y la herramienta MENINA pretenden mejorar el proceso de coexistencia entre IPv4 e IPv6.

Referencias

[1] S. Deering, R. Hinden. "Internet Protocol Version 6 (IPv6) Specification". IETF RFC 2460, Diciembre 1998.

[2] D. Waddington, F. Chang. "Realizing the Transition to IPv6". IEEE Communications Magazine, pp. 138-148, Junio 2002.

[3] M. Tatipamula, P. Grossetete, H. Esaki. "IPv6 Integration and Coexistence Strategies for Next-Generation Networks". IEEE Communications Magazine, pp. 88-96, Enero 2004.

[4] W. Biemolt, A. Durand, D. Finkerson, A. Hazeltine, M. Kaat, T. Larder, R. van der Pol, Y. Sekiya, H. Steenman, G. Tsirtsis. "An overview of the introduction of IPv6 in the Internet". draft-ietf-ngtransintroduction-to-ipv6-transition-09.txt, IETF, Octubre 2002.

[5] M. Lind, V. Ksinant, D. Park, A. Baudot, P. Savola. "Scenarios and Analysis for Introducing IPv6 into ISP Networks". IETF RFC 4029, Marzo 2005.

[6] C. Huitema, R. Austein, S. Satapati, R. van der Pol. "Unmanaged Networks IPv6 Transition Scenarios". IETF RFC 3750, Abril 2004.

[7] Y. Pouffary, J. Bound, M. Blanchet, T. Hain, P. Gilbert, M. Wasserman, J. Goldschmidt, A. Isaac, T. Chown, J. Palet, F. Templin, R. Brabson. "IPv6 Enterprise Network Scenarios". IETF RFC 4057, Junio 2005.

[8] A. Durain, K. El-Maki, N. Murphy, H. Shieh, J. Soininen, H. Soliman, M. Wasserman, J. Wiljakka. "Analysis on IPv6 Transition in 3GPP Networks". draft-ietf-v6ops-3gpp-analysis-11.txt, IETF, Octubre 2004.

[9] T. Miguel, E. Castro. "Programming Guidelines on transition to IPv6". Technical report, IPv6 Forum, Enero 2003.

[10] Inc. Sun Microsystems. "Porting Networking Applications to the IPv6 APIs, Solaris Version 8". Octubre 1999.

[11] M-K. Shin, Y-G. Hong, J. Hagino, P. Savola, E. M. Castro. "Application Aspects of IPv6 Transition ". IETF RFC 4038, Marzo 2005.

[12] K. Tsuchiya, H. Huguchi, Y. Atarashi. "Dual Stack Hosts using the Bump-In-the-Stack Technique (BIS)". IETF RFC 2767, Febrero 2000.

[13] S. Lee, M. Shin, Y. Kim, E. Nordmark, A. Durand. "Dual Stack Hosts using Bump-in-the-API (BIA)". IETF RFC 3338, Octubre 2002.

[14] E. Castro. "Contribución al estudio y definición de una metodología de transición gradual de IPv4 a IPv6". Tesis doctoral, ETSIT-UPM, Septiembre 2004.

[15] J. Arkko, M. Blanchet, S. Chakrabarti, A. Durand, G. Gastaud, J. Hagino, A. Inoue, M. Ishiyama, J. Loughney, R. Raghunarayan, S. Sakane, D. Thaler, J. Wiljakka. "IPv6 Node Requirements". draft-ietf-ipv6-node-requirements-11, IETF, Agosto 2004.

[16] R. Gilligan, E. Nordmark. "Transition Mechanisms for IPv6 Hosts and Routers". IETF RFC 2893, Agosto 2000.

[17] A. Durand, P. Fasano, I. Guardini, D. Lento. "IPv6 Tunnel Broker". IETF RFC 3053, Febrero 2001.

[18] B. Carpenter, K. Moore. "Connection of IPv6 Domains via IPv4 Clouds". IETF RFC 3056, Febrero 2001.

[19] J. Bound, L. Toutain, O. Medina, F. Dupond, M. Shin, J. Lee, H. Lee, E. Castro. "Dual Stack Transition Mechanism (DSTM)". draft-bound-dstm-exp-02.txt, IETF, Noviembre 2004.

[20] C. Huitema. "Teredo: Tunneling IPv6 over UDP through NATs". draft-huitema-v6ops-teredo-05, IETF, Abril 2005.

[21] F. Templin, T. Gleeson, M. Talwar, D. Thaler. "Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)". draft-ietf-ngtrans-isatap-24, IETF, Enero 2005.

[22] T. Bolognesi, E. Brinksma. "Introduction to the ISO specification language LOTOS". Computer Networks and ISDN Systems, 14, 1987.

Jp2p: una infraestructura descentralizada para juegos en red

S. Machado, J.M. Yúfera, X. Barrera
Departamento de Ingeniería Telemática. Universidad Politécnica de Cataluña
Escola Politécnica Superior de Castelldefels, Avda Canal Olímpic, s/n
08860 Castelldefels (Barcelona)
Teléfono: 934 13 70 50 Fax: 934 13 70 07
E-mail: yufera@mat.upc.es

Abstract*. This paper presents Jp2p, an application level scalable infrastructure for network gaming. It supports a high number of groups and of players in each group. The infrastructure is build upon Pastry, a generic peer-to-peer routing and location layer, running upon Internet. The reliability, auto organization and locality properties are inherited by the gaming application from Pastry which, furthermore, will be used for building and organizing the players groups. New methods and messages have been implemented for this infrastructure, which will be used by the network nodes for managing the inputs and outputs from every available match. The designed protocols and obtained results will be explained.*

1 Introducción

Los juegos en red multijugador tradicionales se basan en arquitecturas cliente-servidor donde la escalabilidad se consigue mediante clusters de servidores.

Aunque estas arquitecturas pueden escalar con el número de jugadores, tienen limitaciones de flexibilidad, y el servidor debe estar sobredimensionado para manejar picos de carga. Además, el modelo cliente-servidor limita la implantación de juegos diseñados por el usuario que no constituyan únicamente extensiones o mods de juegos ya operativos.

Sin embargo, debemos pensar que los juegos en red multijugador son aplicaciones adecuadas para las redes *overlay peer-to-peer*, ya que pueden beneficiarse de sus características de auto organización y escalabilidad.

Existen diversos trabajos recientes sobre redes *overlay peer-to-peer* que presentan substratos escalables, capaces de organizarse automáticamente y tolerantes a fallos para aplicaciones distribuidas descentralizadas. Ejemplo de estos pueden ser, entre otros, Chord, Tapestry, CAN o Pastry [1-4].

Basándonos en estas ideas, en este artículo presentamos una infraestructura descentralizada a nivel de aplicación construida sobre Pastry, un substrato *peer-to-peer* para localización y encaminamiento escalable y con capacidad de auto organizarse que presenta buenas propiedades de localidad, motivo por el cual ha sido elegido de entre los candidatos posibles. La infraestructura es escalable a gran número de grupos y de miembros por grupo.

Juntos, la infraestructura para juegos Jp2p y Pastry adoptan un modelo *peer-to-peer* completamente descentralizado en el que cada uno de los participantes tiene las mismas responsabilidades. Sin embargo, es posible la elección de un nodo, de entre los que forman un grupo, para que cumpla funciones de sincronismo de juego.

El enfoque para el caso de juegos en red es diferente del visto en aplicaciones *peer-to-peer* previas, que se centran básicamente en la distribución de contenidos, mensajería instantánea o utilización de recursos. Nosotros dirigimos nuestro trabajo, entre otros puntos, a evaluar la escalabilidad de la infraestructura de juego, su robustez frente a caídas, la formación y gestión de grupos de jugadores o el funcionamiento de la infraestructura según los mecanismos de sincronismo de juego utilizados.

El resto del artículo se organiza del modo siguiente. En la sección 2 se ofrece una idea de la infraestructura de localización y encaminamiento llamada Pastry. En la 3 se explicará el diseño básico de la infraestructura de juegos Jp2p que funcionará sobre Pastry. Después, se presentarán los resultados obtenidos de la simulación de la infraestructura conjunta en la sección 4 para, finalmente, describir las conclusiones obtenidas en la 5.

Las contribuciones técnicas principales de este trabajo son de carácter de evaluación y arquitectura. Se presenta una arquitectura nueva que compagina tecnologías de red *peer-to-peer* con juegos multijugador, un protocolo sencillo para la gestión de grupos y, finalmente, un estudio de las prestaciones de la infraestructura para demostrar la viabilidad de la idea.

2 Pastry

En esta sección se realiza un breve presentación de Pastry [4], el substrato de localización y encaminamiento *peer-to-peer* sobre el que se ha construido la infraestructura de juego Jp2p.

Pastry forma una red *overlay* robusta y capaz de auto organizarse en Internet, en la que puede participar cualquier nodo que utilice el software Pastry y tenga las credenciales adecuadas.

Cada nodo Pastry tiene un identificador único de 128 bits (NodeId) obtenido, por ejemplo, a partir del resultado de aplicar una función de *hash* a su dirección IP, de modo que el conjunto de identificadores de nodos existentes en un espacio circular definido entre 0 y 2^{128}-1 esté uniformemente distribuido.

Dado un mensaje y una clave, Pastry encamina el mensaje hasta el nodo con el NodeId numéricamente más cercano a la clave, de entre todos los nodos Pastry activos en la red, en menos de $\lfloor \log_{2^b} N \rfloor$ saltos en media (donde *b* es un parámetro de configuración con valor típico 4, y *N* es el número de nodos de la red Pastry).

Las tablas requeridas en cada uno de los nodos Pastry para el correcto funcionamiento de la red *overlay* tienen únicamente $(2^b - 1) * \lfloor \log_{2^b} N \rfloor + l$ entradas, donde cada entrada mapea los NodeId a las correspondientes direcciones IP (y donde *l* es un número par, parámetro de configuración de valor típico 16).

Para el encaminamiento de los mensajes, los NodeId y las claves se utilizan como una secuencia de dígitos en base 2^b, y cada una de las 2^b*-1* entradas en la fila *n* de la tabla de encaminamiento se refiere a un nodo cuyo NodeId concuerda con el del actual en los primeros *n* dígitos, pero cuyo dígito *n+1* tiene uno de los otros 2^b*-1* posibles valores. De este modo, y en cada uno de los saltos del encaminamiento, cada nodo envía el mensaje a un nodo de su tabla cuyo NodeId comparta con la clave un prefijo de, al menos, un dígito (o b bits) más largo que el prefijo que la clave comparte con el NodeId del nodo actual (ver Fig. 1).

2.1 Propiedades de localidad

Las propiedades de localidad de Pastry se basan en la métrica de proximidad, un escalar que refleja una idea de distancia entre cualquier par de nodos (como podría ser, por ejemplo, el RTT). Se asume que existe una función que permite determinar a cada nodo su distancia con otro nodo con dirección IP dada.

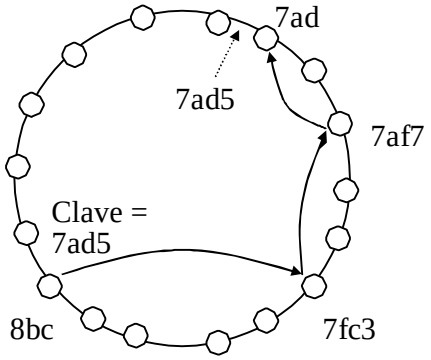


Fig. 1: Encaminamiento del mensaje con clave 7ad5 en el espacio Pastry de identificadores de nodos, desde el NodeId=8bca hasta el 7ad2.

Una de las propiedades de localidad de Pastry es la de rutas cortas, que tiene que ver con la distancia total, en términos de métrica de proximidad, que viajan los mensajes a lo largo de las rutas Pastry.

Como cada entrada en las tablas de encaminamiento se escoge según dicha métrica (el nodo más cercano de entre los posibles según prefijo), en cada paso del encaminamiento el mensaje se envía al nodo más próximo con coincidencia de prefijo mayor. De hecho, la distancia esperada recorrida por el mensaje en el último paso de encaminamiento suele dominar la distancia total recorrida por el mismo. Con estas premisas, y según simulaciones realizadas [5], la distancia viajada por un mensaje en la red Pastry es entre 1,59 y 2,22 veces superior a la distancia entre fuente y destino en la red sobre la que funciona.

2.2 Inclusión y fallo de un nodo

Pastry es eficiente en el mantenimiento dinámico del estado de los nodos, es decir, de las tablas de encaminamiento, en presencia de fallo o llegada de nodos.

Para conseguirlo utiliza tablas de nodos vecinos, entre los que se intercambian mensajes de *keep-alive* periódicamente, y el *leaf set* (un conjunto de nodos numéricamente cercanos a nuestro NodeId) para actualizar las tablas. Como se explicará más adelante, en la infraestructura de juegos Jp2p que funciona sobre Pastry se utilizarán las tablas de vecinos y el *leaf set* para gestionar las posibilidades de juego de cada uno de los nodos.

Para conseguir ser capaz de auto organizarse, Pastry debe mantener dinámicamente el estado de sus nodos (es decir, la tabla de encaminamiento y el *leaf set*) en presencia de la llegada de nuevos nodos, fallos o particiones de red.

2.3 API

La API exportada por Pastry para las aplicaciones como la que se quiere probar aquí incluye, entre otras, las siguientes operaciones:

- *NodeId=pastryInit(Credentials)* que permite a un nodo local unirse a una red Pastry e inicializarse, y devuelve el NodeId del nodo local. Las *Credentials* las ofrece la aplicación y contienen información para la autenticación y seguridad del nodo en la red;
- *Route(msg, key)* permite al nodo local encaminar un mensaje dado hacia el siguiente nodo Pastry, según la clave;
- *Send(msg,IP-addr)* permite al nodo local enviar el mensaje hacia el nodo con la dirección IP especificada.

3 Infraestructura de Juegos

3.1 Idea general

La idea funcional de la infraestructura es la de permitir que diferentes usuarios puedan jugar en red (de manera descentralizada, escalable y con capacidad para auto organizarse), partidas multijugador. Además, los propios jugadores deberían poder publicar de manera sencilla en la red sus propios juegos y hacerlos disponibles para los demás usuarios.

Cualquier nodo que quiera jugar debe poder unirse a un grupo o dejarlo al finalizar la correspondiente partida. Del mismo modo, cada uno de estos grupos tendrá un número de jugadores variable, en función de las posibilidades del juego.

En este contexto, la infraestructura se basa en utilizar un identificador del juego al que el usuario desea jugar, como parte del identificador del nodo local o NodeId. Por esto, Jp2p ofrece una API a sus aplicaciones (en este caso los juegos) que se compone básicamente de:

- *Create (credentials, gameId)* cuando un usuario ha creado un nuevo juego y quiere publicitarlo en la red Pastry;
- *Join (credentials, gameId)* cuando un usuario quiere unirse a algún grupo para jugar;
- *Leave (credentials, gameId)* cuando un usuario decide abandonar una partida;
- *State (credentials, gameId, message)* cuando el juego debe enviar, como refresco, información del estado de su partida a los nodos de la red.

La infraestructura total resultante de unir Jp2p y Pastry es totalmente descentralizada, ya que todas las decisiones se toman en base a información local y, en un principio, todos los nodos tienen idénticas funcionalidades. Esto no significa que no sea posible, también según la información de la red *overlay* con la que trabaja cada uno de los nodos, elegir nodos que realicen funciones de sincronismo de juego.

3.2 Implementación

Una infraestructura completa de juegos debe estar compuesta de un conjunto de nodos Pastry, cada uno de los cuales utilice Jp2p. El programa Jp2p ofrece los métodos *forward* y *deliver*, invocados por Pastry cuando llega a un nodo uno de los mensajes Jp2p. Un ejemplo en pseudocódigo de dichos métodos se muestra en la Fig.2.

Se llama al método *forward* siempre que un mensaje Jp2p se encamina a través del nodo. Por otro lado, al método *deliver* sólo cuando un mensaje alcanza el nodo destino, según el correspondiente identificador y clave Pastry.

Los mensajes Jp2p que se han definido son los siguientes:

- GAMESEARCH: es un mensaje que contiene un vector (*MessageVector*) en el que se almacenan los *nodeId* de los nodos que están esperando para jugar posibles partidas;
- GAMEREQUEST: este mensaje es una petición para unirse a una partida;
- GAMERESPONSE: es la respuesta a GameRequest y le indica a un usuario si ha sido aceptado en la partida o no.
- GAMESTART: este mensaje se envía a los jugadores de una partida para indicarles que pueden empezar a jugar.
- GAMEUPDATE: este mensaje se envía para informar a los jugadores de una partida de que un nodo se ha añadido o se ha dado de baja.
- GAMEOVER: cuando un nodo acaba de jugar, se envía un mensaje GameOver para cada jugador.

Hemos considerado que cada uno de los nodos Jp2p puede encontrarse en uno de los siguientes estados:

- Buscando: cuando el nodo se encuentra buscando una partida a la que unirse;
- Esperando: cuando se ha añadido a una partida, pero ésta no está completa y no se puede iniciar;
- Jugando: cuando el nodo, una vez completado el número máximo de jugadores necesarios, se ha incorporado al juego.

3.3 Gestión del grupo de juego

Cada grupo tendrá en común el juego en el que está participando, de manera que cada jugador deberá decidir a qué quiere jugar para poder calcular su NodeId, que se obtendrá de encadenar el resultado de una función de *hash* del nombre del juego y el de un identificador del nodo (como por ejemplo, su dirección IP).

```
forward (msg, key, nextId)
   switch msg type is
      GameSearch: if nodo esta ESPERANDO
                        GameSearch.Vector∪=NodeId
                        route(nexthop,GameSearch)

deliver (msg, key)
   switch msg type is
      GameRequest:
            if nodo esta ESPERANDO
                  if partidaCompleta
                        confirm=false
                        GameResponse(confirm,)
            else
                        confirm=true
                        GameResponse(confirm, partida)
            for resto jugadores send(jugador,GameUpdate)
            send(GameRequest.origen, GameResponse)

      GameResponse:
            if GameResponse.confirm
                  estado=ESPERANDO
                  app.partida=msg.getPartida()
                  if partidaCompleta
                        for cada jugador
                           send(jugador,GameStart)
                  else
                        posiblepartida = vector.next()
                        if posiblepartida=null
                              send(nexthop,GameSearch)
                        else
                              send(posiblepartida, GameRequest)

      GameStart:
            if msg.origen==localNode
            app.jugar

      GameSearch:
            if nodo=GameSearch.origen
                  vector=GameSearch.vector
                  posiblepartida = vector.next()
                  if posiblepartida=null
                        send(GameSearch, nexthop)
                  else
                        send(GameRequest, posiblepartida)
                  else
                        if nodo esta ESPERANDO
                           GameSearch.vector∪=NodeId
                     send(GameSearch.origen,
GameSearch)

      GameUpdate:
            if msg.esNodoNuevo
                  partida.agregarJugador(msg.getJugador)
            else
                  partida.borrarJugador(msg.getJugador)

      GameOver:
            partida.acabados++
            if !msg.seguirJugando
                  partida.borrarJugador(msg.getJugador)
```

Fig.2: Pseudocódigo de los métodos Jp2p invocados por Pastry.

Una vez elegido el juego en el que un usuario desea participar, éste debe unirse a la red Pastry mediante un mensaje Pastry de tipo JoinRequest. El usuario pondrá de destinatario/clave su propio NodeId, de esta forma Pastry encaminará el mensaje hacia el destinatario con el NodeId más parecido al NodeId del nuevo usuario. Esta funcionalidad de Pastry sólo se consigue siempre y cuando el mensaje original vaya encapsulado en un mensaje Pastry de tipo RouteMessage.

El RouteMessage recorrerá la red de nodos Pastry hasta que encuentra el nodo con el NodeId más cercano a la clave. Este nodo extraerá el contenido del RouteMessage y se encontrará con el JoinRequest. El nodo reenviará al origen el JoinRequest aceptándole el Join.

Jp2p aprovecha el mecanismo de Join de Pastry para buscar partidas incompletas del juego elegido por el nuevo usuario. Esto se consigue gracias a que el JoinRequest contiene un vector donde se almacenan las partidas que se han ido encontrando durante el encaminamiento hacia el destino.

A cada paso del encaminamiento del RouteMessage se mira si el nodo intermedio pertenece al mismo grupo y si está en estado Esperando. Si esto se cumple, el nodo intermedio se añade al vector de partidas incompletas.

Cuando el JoinRequest regresa al nuevo usuario, éste extrae el vector de partidas incompletas y elige una a la que se quiere unir. Para ello le envía al nodo intermedio que está en esa partida, un mensaje Jp2p de tipo GameRequest.

Con el mensaje GameRequest le indicamos al nodo que está Esperando que queremos unirnos a la partida. Aquí pueden pasar dos cosas:

- La partida aún está incompleta: en ese caso el nodo que está Esperando añade el nuevo usuario a la partida, envía un mensaje GameUpdate al resto de jugadores para indicarles que se ha añadido un nodo, y le envía un mensaje Jp2p de tipo GameResponse para comunicarle al nodo nuevo que ha sido aceptado. GameResponse contiene un parámetro denominado 'confirm' que indica si el nodo ha sido aceptado en la partida o no; en caso de que sí fuera aceptado, el mensaje también contendría la partida. Una vez el nodo reciba el GameResponse, mira el parámetro 'confirm', y si ha sido aceptado se pone en estado Esperando. Si el nodo es aceptado en la partida y resulta que ésta se completa, la partida envía a cada jugador un mensaje GameStart avisando a los nodos de que pueden empezar a jugar. En ese momento, los estados de los nodos jugadores pasan a Jugando.

- La partida está completa: en este caso el nodo le envía un mensaje Jp2p de tipo GameResponse con el parámetro 'confirm' igual a false. El nuevo nodo recibirá el GameResponse y borrará esa partida del vector de partidas incompletas, para luego escoger una nueva partida del vector. Si se da el caso de que el vector de posibles partidas se queda vacío, se buscarán nuevas partidas mediante un mensaje GameSearch (más adelante explicaremos el mecanismo de este mensaje).

Cabe destacar que los mensajes Jp2p no pueden encaminase por la red Pastry por sí solos, sino que deben encapsularse dentro de un mensaje Pastry de tipo RouteMessage. Pero en Jp2p sólo se encaminan los mensajes GameSearch, el resto se envían directamente hacia el nodo destino.

Cuando un jugador quiere dejar de jugar o finaliza la partida, tiene tres opciones:

- Puede quedarse en la misma partida esperando a que se complete otra vez el número de jugadores para volver a jugar. En este caso el estado del nodo pasa de Jugando a Esperando
- Puede cambiar de partida e incluso de juego. Para ello se hace uso del mensaje Jp2p de tipo GameSearch y en ese caso su estado pasa a ser Buscando.
- Otra opción sería que el jugador ya no quisiera jugar más. En ese caso el jugador se borra de la partida y el nodo es eliminado de la red. Al igual que pasa cuando se quiere cambiar de partida, si un jugador se borra de una partida y en ésta ya no quedan más jugadores, la partida es eliminada.

Cuando un nodo acaba de jugar, debe avisar al resto de los jugadores de la partida con un mensaje GameOver.

En el segundo caso, si un jugador quiere cambiar de partida, antes de todo tiene que buscar partidas incompletas por la red Pastry. Esto se hace de una forma similar al JoinRequest.

En primer lugar el nodo tiene que construir una ruta por donde pasará el mensaje GameSearch. Para ello hace uso de su *leafset* y su tabla de ruta. La ruta se compone simplemente de un vector que contiene todos los NodeIds de los nodos que hay en el *leafset* y la tabla de ruta (no incluyéndose él mismo).

El vector con la ruta se almacena en el mensaje GameSearch y se envía al primer nodo del vector. A parte de este vector, el GameSearch contiene un vector de partidas incompletas donde se van almacenando los NodeIds de aquellos nodos que están en una partida incompleta. Cuando el mensaje GameSearch llega al primer nodo, se mira si éste está en estado Esperando. Ésta es la única condición para añadir el NodeId del nodo al vector de partidas incompletas. Luego el mensaje se envía al segundo nodo, y así sucesivamente. Finalmente, cuando ya ha recorrido toda la ruta, el mensaje GameSearch regresa al nodo origen, y el usuario elige una partida del vector de partidas incompletas.

Si la partida elegida es de otro juego, el nodo necesita cambiar de NodeId. El cambio se hace antes de enviar el mensaje Jp2p GameRequest. Por lo tanto, lo único que diferencia un nodo de un grupo o de otro es su NodeId.

4 Resultados Experimentales

Para comprobar el funcionamiento de la infraestructura completa de juegos, se programó sobre FreePastry, una implementación Java del Pastry realizada por la Rice University [6]. FreePastry emula una red Pastry con un gran número de nodos. Las pruebas realizadas han servido para evaluar el Jp2p frente a: retardo entre los miembros de los grupos formados, caída de nodos,...

La aplicación programada permite probar y observar, mediante su interficie, algunas funciones de Pastry, como son: el envío de mensajes, el número de saltos en el encaminamiento, la creación de nodos, el comportamiento cuando se borra un nodo, los cambios en el *leafset* y la tabla de ruta.

También se pueden controlar aspectos de la infraestructura Jp2p, como a qué juego jugará cada nodo, elegir a qué partida se deben unir, y controlar las decisiones en el fin de partidas.

Además, gracias a que FreePastry emula la red, la interficie de la aplicación programada muestra en todo momento la topología de la red, las conexiones entre los nodos, e incluso el estado de éstos. De esta forma se pueden observar en todo momento los cambios en los nodos durante las pruebas y analizar su comportamiento.

Para las pruebas se consideraron 3 juegos en la red con capacidad para 4 jugadores. Los 3 juegos son exactamente idénticos, lo único que les diferencia es el nombre del juego, con el que un nodo puede identificar a uno u otro. Cada juego contiene solamente un contador aleatorio de tiempo de juego, que simula el comportamiento de un jugador real en la red.

Para llevar a cabo las pruebas, la aplicación puede crear una red con un número aleatorio de nodos, pudiendo escoger de cada nodo el juego deseado también de manera, además de las partidas a las que quiere jugar. El tiempo de juego también es aleatorio, y los nodos creados escogen una de las opciones de fin de juego aleatoriamente. De esta forma con un simple clic a un botón, la aplicación nos crea una red de nodos Jp2p totalmente aleatoria que sirve para obtener resultados del tipo: nodos creados, nodos eliminados, partidas creadas, partidas eliminadas, tiempo medio en que un jugador está esperando, … Además mediante la interficie de la aplicación observamos todos los cambios que van sucediendo en los nodos en tiempo real.

Los resultados que se exponen a continuación se han obtenido a partir de diez pruebas totalmente aleatorias en redes simuladas de hasta 100 nodos.

Como cada nodo escoge el juego aleatoriamente, las partidas creadas de un juego o de otro se crean con la misma probabilidad.

En el gráfico (Fig.3) se puede observar que la evolución del número de partidas creadas tiene dos fases: una transitoria, en la que cuantos más nodos hay en la red más partidas se crean; y otra de estabilidad. De hecho, a partir de los 20 nodos, la gráfica crece más lentamente. Esto es debido a que cuando hay pocos nodos en la red, se crean más partidas de las necesarias a causa de que hay pocos nodos de un mismo grupo.

Por otro lado, recordemos que una partida es eliminada cuando no contiene ningún jugador. Durante las 10 pruebas realizadas, sólo fue eliminada una partida (Fig.4). Esto es debido a que es muy difícil que una partida se encuentre sin jugadores, ya que los jugadores entran y salen de las partidas constantemente.

En la tabla 1 se muestra la probabilidad de que un nodo acceda a una partida realizando uno o más intentos, según el tamaño de la red simulada. Como se observa en la tabla, la mayor parte de los nodos es capaz de encontrar una partida y acceder a ella en uno o dos intentos.

Otro resultado que se obtuvo fue el tiempo medio en el que un nodo está en estado de espera. En la figura 5 se observa que cuantos más nodos hay en la red, el tiempo de espera es mayor. Este resultado puede parecer incoherente debido a que en las redes con un número elevado de nodos, encontrar jugadores que completen las partidas es más fácil que en las redes poco pobladas. Sin embargo, también debemos tener en cuenta que cuantos más nodos se está emulando, más sobrecargada está la aplicación.

5 Conclusiones

En este artículo hemos presentado una nueva infraestructura escalable para juegos en red. Se ha explicado su funcionamiento y se han mostrado las primeras pruebas realizadas.

Se ha demostrado la viabilidad de la utilización de Pastry como base de la infraestructura de juegos que hemos denominado JP2P así como el buen funcionamiento de los protocolos implementados a la hora de encontrar partidas de juego accesibles prácticamente al primer intento.

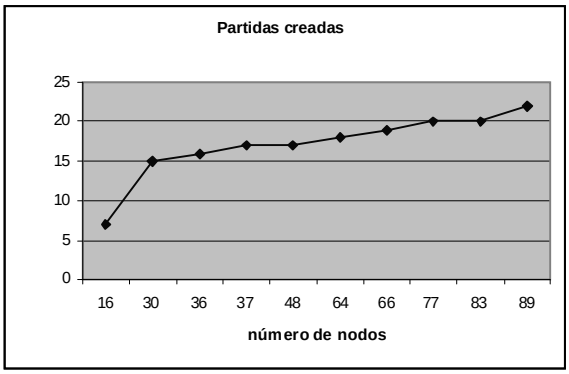


Fig. 3 Gráfico que representa las partidas creadas respecto al número de nodos.

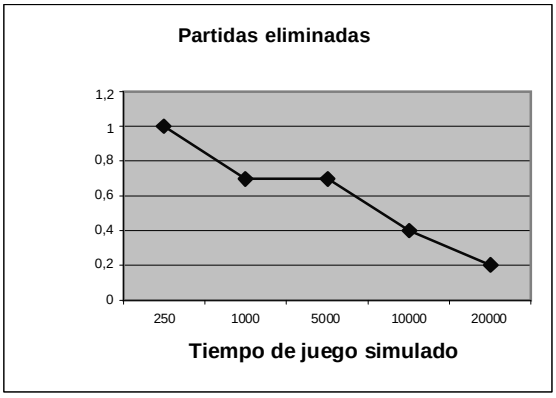


Fig. 4 Gráfico que representa la media de partidas eliminada respecto el tiempo de juego de los usuarios.

Tabla 1 Probabilidad de acceder a una partida en diferentes intentos.

	5 nodos	10 nodos	20 nodos	30 nodos	50 nodos
1er intento	93,4%	93,26%	80,14%	90%	77,3%
2º intento	4,2%	4,78%	8,85%	6,94%	12,27%
3er intento	2,4%	1,96%	7,4%	1,96%	5,07%
4º intento	0%	0%	2,91%	0,55%	2,68%
5º intento	0%	0%	0,7%	0,55%	2,09%
6º o más intentos	0%	0%	0%	0%	0,59%

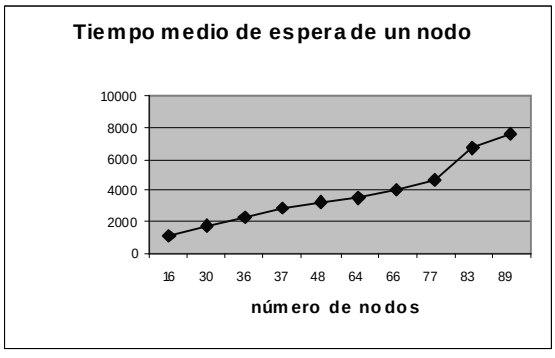


Fig. 5 Gráfico que representa el tiempo medio en el que un nodo está en el estado Esperando, respecto al número de nodos en la red.

Referencias

[1] I.Stoica, R.Morris, D.Liben-Lowell, D.R.Kargen, M.F.Kaashoek, F.Dabek and H.Balakrishnan, "Chord: a scalable peer-to-peer lookup protocol for Internet applications", IEE/ACM Transactions on networking, Vol.11, No.1, 2003, pp.17-32.

[2] B.Y.Zhao, L.Huang, J.Stribling, S.C.Rhea, A.D.Joseph and J.D.Kubiatowicz, "Tapestry: a resilient global-scale overlay for service deployment", IEEE Journal on selected areas in communications, Vol.22, No.1, 2004, pp.41-53.

[3] S.Ratnasamy, P.Francis, M.Handley, R.Karp and S.Shenker, "Application-level multicast using content-addressable networks", Proceedings of the Third International workshop on networked group communication, UCL, London, November, 2001.

[4] A.Rowstron and P.Druschel, "Pastry: scalable, distributed object location and routing for large-scale peer-to-peer systems", International conference on distributed systems platforms (middleware), Nov. 2001.

[5] M.Castro, P.Druschel, Y.Charlie Hu and A.Rowstron, "Exploiting network proximity in peer-to-peer overlay networks", Technical Report MSR-T-2002-82, 2002.

[6] "Pastry-a scalable, decentralized, self-organizing and fault-tolerant substrate for peer-to-peer applications": http:// http://freepastry.rice.edu/

Integración de Servicios Multimedia en Redes 4G

Raúl Sánchez Martín, Antonio Cuevas Casado, Jose Ignacio Moreno Novella, Pedro A. Vico Solano
Departamento de Ingeniería Telemática. Universidad Carlos III de Madrid
Avda. Universidad 30.
28911 – Leganés (Madrid)
E-mail: {rul_l;acuevas;jmoreno;pvico}@it.uc3m.es

Abstract*. This paper presents a solution to integrate multimedia services with a 4G networks which thus becomes a service aggregator and not a mere bits transporter. The solution focuses on SIP services and requires no modification to the SIP UA, only to the SIP-Proxy. Our solution has been implemented and tested in a prototype 4G network "Moby Dick". This serves to demonstrate our concepts and shows the advantage of employing applications being capable of interacting with the network operator infrastructure*.

1 Introducción

En las redes 4G las figuras del proveedor de aplicaciones y proveedor de red se separan, viendo estos últimos el peligro que conllevaría convertirse en meros "transportistas de bits". Para evitarlo han surgido iniciativas que permiten al proveedor de red obtener más protagonismo en la provisión de servicios de valor añadido, por ejemplo, convirtiéndolo en un agregador de servicios. Actualmente estos servicios son básicamente provisión de contenidos; en cambio, las soluciones para los servicios de conferencia están prácticamente sin abordar (excepción hecha del IMS [5] definido por el 3GPP). Además, las soluciones propuestas se centran en redes 3G y no en redes 4G. Este artículo pretende cubrir este hueco y propone e implementa un modelo de integración de servicios de conferencia con proveedores de red 4G. Aunque nos centramos en servicios de video telefonía IP establecidos mediante SIP, nuestra solución podrá ser extensible a escenarios similares y otras aplicaciones con el único requisito de utilizar el protocolo SIP y un SIP proxy para el establecimiento de sesiones.

Aún no existe una definición clara de las redes de la nueva generación, pero si muchos proyectos y prototipos en vías de desarrollo. Moby Dick [4] es uno de ellos; establece una red 4G IPv6 integrada con cualquier tipo de red de acceso (802.11, TD-CDMA, Ethernet) con características de movilidad, AAA y QoS. Además, está implementada y disponible para hacer pruebas. Como en este artículo no sólo diseñamos sino que queremos implementar y mostrar de forma tangible la viabilidad de nuestra solución, elegimos Moby Dick como prototipo de red 4G. Pero, conceptualmente, podría haber sido reemplazada por cualquier otro tipo de red 4G.

Para ver cómo hemos logrado nuestro objetivo dividiremos el artículo en dos partes: primero describiremos el diseño de nuestra solución y luego veremos cómo se ha probado e integrado en un escenario de red 4G. Por último, las conclusiones recogen los principales puntos y plantean posibles trabajos futuros.

2 Descripción de la solución propuesta

2.1 Escenario de red, componentes e información intercambiada

Al apoyarnos en Moby Dick utilizamos las entidades de las que se compone esta red. Para el caso que nos ocupa, conviene clasificarlas según pertenezcan a una u otra de las dos características de esta red que más nos interesan: QoS y AAA. Los aspectos de movilidad disponibles en Moby Dick no han sido aprovechados y quedan para trabajos futuros.

Las entidades que intervienen en la provisión de la QoS (Quality of Service) [10] utilizan el protocolo COPS [3] sobre IPv6 para comunicarse entre si. Estas son:

Los routers que componen la red. Pueden diferenciarse en dos tipos: 1) Los Access Router (AR), se encuentran al borde de la red y se encargan de unir a los usuarios a la misma. 2) Los routers de núcleo, se encuentran en el interior de la red y se limitan a rutar mensajes priorizando unos sobre otros siguiendo las órdenes del QoS Broker y mirando el campo DSCP. También están los routers de borde que conectan el dominio con Internet.

QoS Manager. Se encuentra instalado en los AR. Informa al QoS Broker de los orígenes y destinos y DSCPs de los flujos (IPv6) que establece cada usuario y recibe las órdenes de configuración de éste para darles la QoS apropiada.

El QoS Broker (QoSB) –también llamado Bandwidth Broker- es la entidad que ordena (mediante COPS) a los routers cómo actuar indicando qué paquetes tienen mayor prioridad. Para ello se basa en los perfiles de usuario recibidos del servidor AAA. Con ese perfil se establece la QoS que disfrutará el

usuario cuando se transporten sus datos, pero sin especializarla para ningún servicio.

Los componentes que forman el sistema AAA (Authentication, Authorization and Accounting) [9] emplean el protocolo DIAMETER [1] (sobre IPv6). Son dos:

AAA Server. Almacena en una base de datos la información de los clientes, sus tarifas y otros datos de interés como pueden ser los códigos DSCP que puede emplear el usuario para cada tipo de tráfico. En caso de "roaming" el AAA Server del dominio local actúa de intermediario con el AAA Server del dominio hogar del usuario. Durante el registro, el AAA Server local transmite –usando COPS- al QoSB local parte del perfil del usuario. Este perfil incluye la CoA (Care of Address) del equipo que está empleando el usuario. Si dicho equipo tiene varios interfaces, será la CoA del interfaz conectado actualmente a la red. La CoA será la dirección origen de todos los paquetes enviados por el Terminal del usuario.

AAA Client. Está instalado en los AR. Obtiene la información del AAA Server. Controla el acceso de los usuarios a la red y establece un túnel IPSec IPv6 entre él y el terminal del usuario que, entre otras cosas, garantiza la correspondencia entre usuario y dirección origen del paquete. Se encarga, además, de medir los bytes y paquetes transmitidos y recibidos por el usuario. Esa información la envía al AAA Server –quien la reenviará al AAA Server hogar del usuario si éste está en roaming- y con ella se elaborará una tarificación basada únicamente en tiempo de conexión a la red y volumen de datos transmitidos y recibidos.

La contribución de este artículo es la integración de dos nuevos componentes en el escenario para establecer las llamadas SIP multimedia de voz y datos y que este servicio se integre con el servicio del operador de red Moby Dick:

User Agent (UA) de SIP. Aplicación que corre en el Terminal del usuario que se encuentra fuera de la red accediendo a ella a través del AR. Se implementa en con la aplicación SIP-Communicator del proyecto NIST-SIP [11]. Dicha aplicación no ha sido modificada.

Proxy Server. Es el proxy que mantiene control sobre los estados y las sesiones de los usuarios que se registran para obtener el servicio de llamadas. Viene implementado en la aplicación Jain-Sip-Presence-Proxy, también de NIST-SIP. Originalmente, simplemente rutaba mensajes SIP para establecer, mantener y liberar sesiones entre los UA. Nuestro trabajo logra que pueda informar de estos eventos (establecimiento y liberación) al AAA Server para que lleve una tarificación de las llamadas y al QoS Broker para que garantice la calidad de servicio.

Recalcamos que nuestra solución no ha necesitado modificar los SIP-UA, sólo los SIP Proxy, las ventajas de este enfoque se discuten en [12].

Comunicación Proxy Server - QoS broker.

El protocolo COPS implementado no sigue al pie de la letra las especificaciones de [3]. En este caso las figuras del PDP y PEP se ven ligeramente modificadas y conviene hablar mejor de servidor y cliente. El QoS Broker (QoSB) es quien actúa de servidor y hasta él llegan la información de nuevas conexiones dentro de la red para que decida qué routers configurar y así garantizar una determinada calidad de servicio para los flujos multimedia. Se apoya principalmente en los mensajes:

Report State (RPT) Transporta los parámetros del flujo multimedia (Dirs. Origen y Destino y DSCP). Lo utiliza el Sip Proxy para notificar inicios y finales de llamada al QoS Broker. Con esta interacción se logra dotar de QoS especializada a los flujos intercambiados por los usuarios una vez establecida la sesión con SIP. De no existir esta integración entre el SIP-Proxy y el sistema de QoS de Moby Dick, dicho servicio multimedia recibiría un tratamiento genérico y no especializado determinado, como hemos visto, cuando el usuario se registra en la red.

Comunicación Proxy Server –AAA Server.

La contribución que hace este proyecto en el ámbito AAA es permitir una tarificación especializada para el servicio que se ofrece y no como en Moby Dick donde sólo se tarifica por conexión a la red y envió y recepción de paquetes. Se consigue utilizando el protocolo DIAMETER, siguiendo las pautas definidas en [8]. El cliente AAA estará representado por el Sip Proxy, mientras que el servidor es el mismo AAA Server de Moby Dick. A cada menaje que envía el cliente, recibe una contestación del servidor, de la forma que se explica a continuación:

Accounting-Request (ACR), contiene el NAI (Network Access Identifier) del llamado o del llamante (acuevas@ipv6.it.uc3m.es, por ejemplo) que es el mismo que su SIP-URI. El SIP Proxy los utiliza para informar al AAA Server de inicios y finalizaciones de llamada.

Accounting-Answer (ACA). Es la respuesta del servidor al ACR

2.2 Secuencia del intercambio de mensajes

Consideramos un escenario en el que dos dominios Moby Dick independientes ofrecen el servicio de llamada SIP multimedia a dos usuarios, uno de cada dominio, que quieren establecer una conversación entre sí. Ambas redes tendrán sus propios AAA Server, QoS Broker y Sip Proxy.

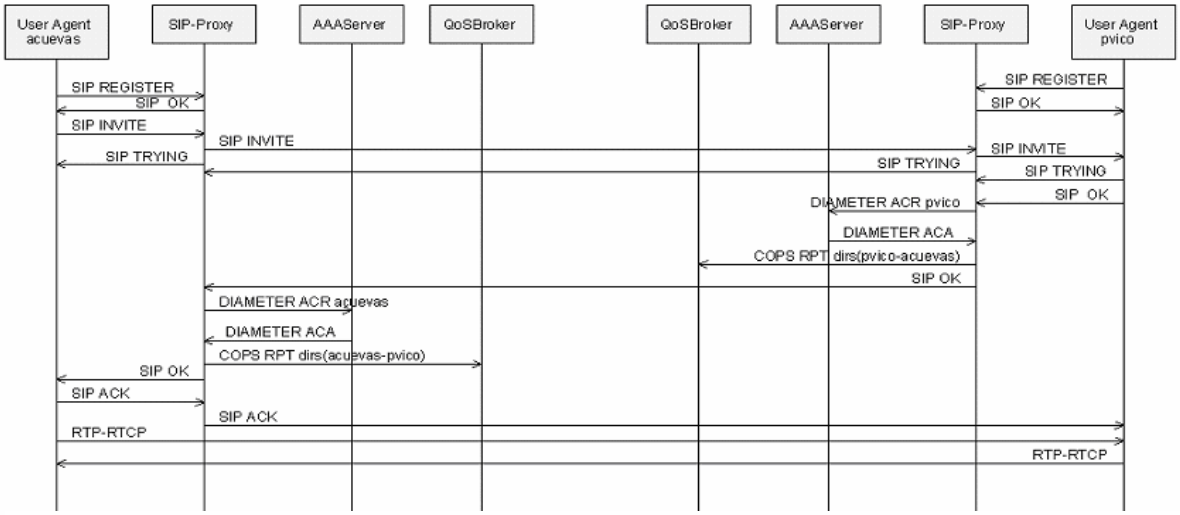


Figura 1 Secuencia de Mensajes para el establecimiento de la llamada

En la Figura 1 se muestra la secuencia de mensajes para el establecimiento de una llamada y la interacción entre los SIP Proxy, los QoS Brokers y los servidores AAA. El lector verá como el SIP Proxy interactúa con el QoS Broker y el AAA Server al recibir el mensaje 200 OK. Varios fueron los motivos para seguir esa aproximación. Primero, en el mensaje 200 OK, ya están presentes todos los parámetros que van a regir la comunicación (CODEC, ancho de banda necesario, …). Además, en el entorno SIP en el que nos basamos [11], ese mensaje recorre todos los SIP Proxy. En cambio el mensaje ACK, al interpretarse extremo a extremo en los clientes, puede ahorrar saltos en el camino. Nótese que esto se podría evitar modificando los Proxies para que rellenaran el campo record-route de la cabecera SIP.

En [4] y en IMS [1] se establece que se inicie el proceso de QoS al recibir el mensaje "183 Session Progress" que contiene los mismos datos (incluido los datos SDP) que contiene el mensaje "200 OK" de respuesta al método INVITE. En IMS, la interacción la lleva acabo el SIP Proxy (CSCF en terminología IMS) a través del PDF con el GGSN (nodo de red 3G) pero sólo para autorizar la futura petición de QoS por parte del usuario. Recalcamos que, según [5], referente a IMS, se da la posibilidad de que esa interacción la realice el CSCF al recibir el mensaje INVITE. En [4], la interacción con QoS la hace el Terminal para reservar QoS. En [12], la interacción la hace el SIP Proxy con los routers y estos interactúan con el Bandwidth Broker (QoS Broker en nuestra terminología). Ésta se hace al recibir el mensaje 200 OK, respuesta al método invite. Como se ve, las aproximaciones son múltiples y como la implementación [11] –que es la que usamos- no emplea el mensaje "Session Progress", nosotros desencadenamos, como hemos dicho, la interacción SIP Proxy-QoS Broker al recibir el mensaje 200 OK. Al centrarse nuestro escenario en redes 4G, esta interacción no se puede hacer con el GGSN nodo de las redes 3G y que, presumiblemente, no estará en las redes 4G. En concreto, no existe en la red 4G –Moby Dick- empleada para validar nuestro trabajo. Por otra parte, decidimos que el SIP Proxy no interactuara y configurara los routers directamente sino que esto se hiciera a través del QoS Broker. Esto tiene una clara ventaja: se soportan los dos modelos antes expuestos -no se sabe aún cual (o cuales) de estos modelos adoptarán las redes 4G-. Por una parte el QoS Broker puede interpretar el mensaje como una autorización y con ella permitir las futuras peticiones de reserva de QoS. La segunda aproximación es que el QoS Broker, al recibir el mensaje, configure directamente los routers. Veremos en la sección 3.1 cual de las dos aproximaciones es la seguida en Moby Dick.

En cuanto a la interacción con el AAA Server, nosotros sólo la hacemos para logar tarificación (Accounting). La autenticación y autorización se dejan para futuros trabajos. En cualquier caso, mencionamos que si un usuario es capaz de enviar mensajes a través de la red y, en particular los mensajes SIP al SIP Proxy, es que éste ya ha sido autenticado y autorizado por el servidor AAA. En IMS esa interacción se hace con el HSS al recibir el mensaje 200 OK, respuesta al mensaje INVITE. El HSS es un nodo de red 3G que en redes 4G será sustituido, con seguridad, por un servidor AAA DIAMETER. Recalcamos que también sería lógico hacer esa interacción al recibir el mensaje ACK.

El intercambio de mensajes implementado cuando un usuario quiere finalizar la conversación se encuentra en la Figura 3, representado en la líneas no punteadas. En [12] entorno SIP en el que nos basamos, los mensajes BYE se interpretan extremo a extremo, es decir, no es necesario que pasen por todos los Sip Proxy que intervienen en la llamada, ni tan siquiera por los dos que se encuentran a los extremos, a los que se conectan los User Agent. Este comportamiento dificulta la liberación de recursos dentro de la red Moby Dick. Será únicamente un Sip Proxy quien verá el mensaje, el conectado al usuario que cierra la conexión.

En Moby Dick no es posible la comunicación entre QoS Brokers de diferentes dominios, siendo la administración de la QoS de forma local. Por este motivo, solamente el primer Sip Proxy podrá liberar los recursos de su red. Este problema no se encuentra a la hora de terminar las tarificaciones de la llamada, pues si se permite que los AAA Server hablen entre sí: el AAA Server local del usuario que finaliza la llamada actuará como Intermediario para el AAA Server del otro interlocutor indicándole tal situación.

Modificando los mensajes SIP (añadiéndose los proxies en el campo record-route de la cabecera SIP) se puede conseguir que el mensaje BYE pase por todos los SIP proxies por los que pasó el mensaje 200 OK, respuesta al mensaje INVITE. Una comprobación similar a la hecha con ese mensaje, viendo si se es el primer o último Sip Proxy bastaría para descartar la necesidad de actuar (comunicar con el AAA y el QoSB) en posibles saltos intermedios. Esta solución, representada en la Figura 3 por las líneas punteadas, sería la óptima pero su implementación aún no está finalizada.

2.3 Soporte de “Roaming”

El protocolo SIP y en especial el Proxy Server empleado, permite el rutado de los mensajes SIP entre Proxies SIP de distintos dominios por lo que el aspecto de roaming no presenta ninguna dificultad añadida. En cuanto a la interacción del Proxy Server con los QoSB y los servidores AAA, ésta siempre se hace en el dominio local (o visitado según la nomenclatura).

En lo que respecta a los QoSBs, como estos actúan a nivel local (configurando los routers de su dominio), el aspecto de “roaming” no supone ningún problema.

Si nos preocupamos por los AAA Servers, el AAA Server del dominio local no tiene el perfil del usuario y por lo tanto no juega ningún papel en este escenario. Pero, el protocolo DIAMETER contempla el rutado de mensajes entre dominios y, en nuestro caso, el AAA Server local actuaría como un enlace con el AAA Server hogar. Por lo tanto la interacción AAA Server – SIP Proxy tampoco presenta ningún problema en el caso de “roaming”.

El único problema se reduce a la forma en que el usuario conocerá la dirección del Sip Proxy de la red a la que se ha conectado (red visitada o red local). Las formas de solucionarlo pueden ser muy variadas incluyendo extensiones a DHCP IPv6 o añadir esta información en los “Router Advertisement” emitidos por los routers de acceso.

Como se puede ver la situación de roaming es

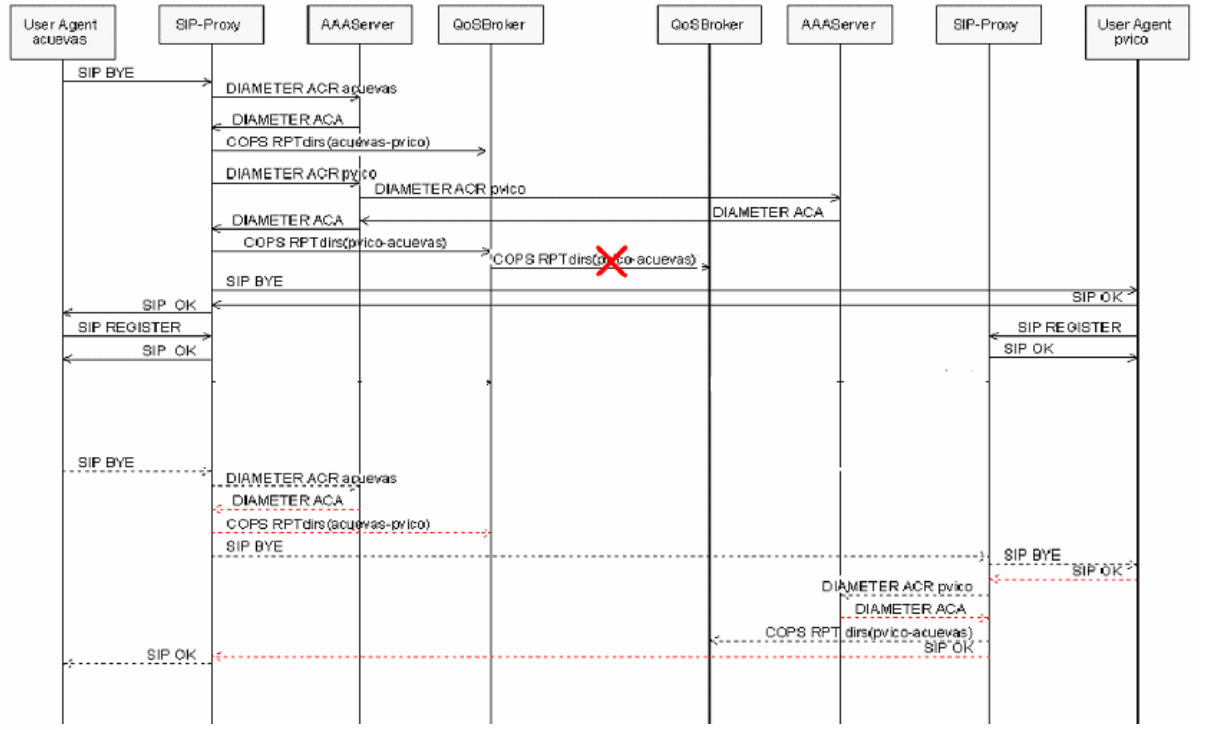


Figura 2 Secuencia de mensajes para finalización de la llamada

perfectamente compatible con el modelo que se propone en este artículo, dejando la responsabilidad a los proveedores de servicios quienes deben acordar las condiciones en las que permitir roaming.

3 Implementación, Integración y Pruebas

3.1 Modificación del Código del SIP Server

De la sección anterior se pueden obtener dos situaciones generales relacionadas con la QoS y AAA, en las que el Sip Proxy debe comunicarse con el AAA Server y el QoS Broker: el inicio y la finalización de las llamadas. En el diagrama de flujo de la Figura 4 quedan reflejadas las comprobaciones que hace, gracias a nuestra implementación [13], el Sip Proxy cuando recibe un mensaje SIP, en base a las cuales sabrá si debe o no actuar (contactar con el QoSB y el AAA Server).

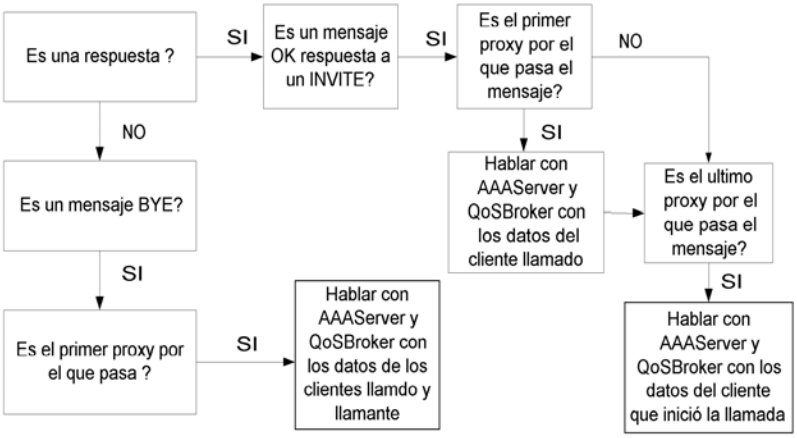


Figura 4 Mecanismo para detectar en el SIP proxy la necesidad de interacción con QoS y AAA

La primera comprobación que se realiza parece bastante trivial, y puede no parecer la más adecuada para detectar las situaciones de inicio y finalización de llamada. En cambio, no es así, debido a que esta diferenciación la hacía el propio Sip Proxy antes de ser modificado. De esta forma se facilita mucho el trabajo ya que un mensaje BYE nunca será una respuesta, mientras que un mensaje 200 OK lo será siempre.

Cuando se recibe una respuesta 200 OK a un mensaje INVITE se aprovecha que estos mensajes se interpretan en todos los Sip Proxies por los que pasó la invitación. En cada uno de ellos se comprueba si se es el primero y/o el último, lo que asegura que son los proxies del dominio local (donde se encuentra conectado el usuario). Sólo aquellos que lo cumplan serán quienes se comuniquen con sus respectivos AAA Server y QoS Broker indicándoles los datos del interlocutor que se encuentra en su red para que inicie su tarificación y reserve los recursos que tiene contratados. En el caso de que ambos interlocutores se encuentren en la misma red, esta operación se hará por duplicado.

Cuando el mensaje es un BYE, sólo atravesará un Sip Proxy, será él quien deba hablar con el QoS Broker y AAA Server locales para finalizar la tarificación y liberar recursos; con la consiguiente problemática que conlleva el no poder liberar los recursos en la red de quien no inicia la finalización de la llamada, como ya se ha explicado anteriormente. Como hemos visto en la sección 2.2, esto es un mero problema de implementación.

3.1 Pruebas en el Escenario Integrado

En la Figura 4 se muestran el escenario de pruebas con todos los elementos y redes físicas, tanto los pertenecientes a la red Moby Dick, como los aportados por nuestro proyecto. El escenario usa exclusivamente IPv6 nativo. Consideramos un único dominio y dos usuarios pertenecientes a ese mismo dominio.

En punteado se muestran los mensajes “puramente Moby Dick”, es decir mensajes en los que este artículo no tiene ninguna participación. Los describimos brevemente: los mensajes entre los dos terminales (chinfano y abeja) y el router de acceso –AR- (termita), son mensajes de registro en la red, que son traducidos por el cliente AAA del AR a mensajes DIAMETER (mensajes entre termita y el AAA Server). El AAA Server además informa del perfil del usuario al QoSB.

Los mensajes entre el QoS Manager de termita y el QoS Broker son para configurar el router de acceso a cada vez que un flujo originado por cualquiera de los dos terminales pasa por él.

Los mensajes obtenidos en este escenario serían los mismos que los descritos en la Figura 1 y la Figura 3, si se unen los dos AAA Server, los dos QoS Broker y los dos Sip Proxy entre sí. Al estar todos los elementos dentro de la misma red Moby Dick y el mismo dominio, no hay que preocuparse de la liberación de llamada, pues el Sip Proxy que se encuentra en abejorro será a la vez el primero y el último tanto en el establecimiento como en la liberación de la llamada. Al detectar ambas situaciones, el SIP proxy hablará tanto con el AAA Server como con el QoS Broker por cada uno de los usuarios utilizando los protocolos DIAMETER y COPS respectivamente.

Los mensajes RTP y RTCP se envían entre los usuarios cuando están en conversación. Por la topología del ejemplo, al pertenecer a una red

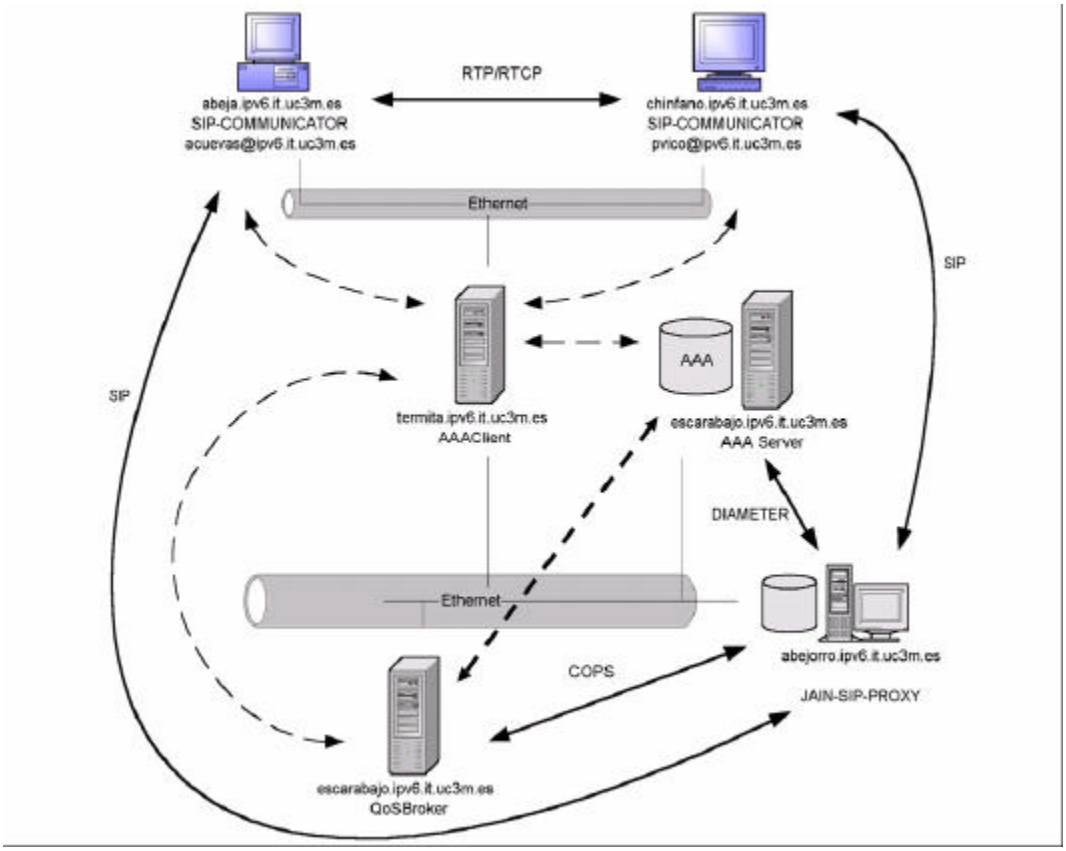


Figura 5 Escenario de Pruebas

Ethernet común pero aislada del resto, estos mensajes no entrarían en la red Moby Dick (no atraviesan el router de acceso). Por este motivo no se les estaría ofreciendo un servicio real de QoS, pero si estuviesen separados se verían obligados a enviar los mensajes a través de la red 4G y disfrutarían de sus ventajas. En concreto, al pasar el flujo RTP/RTCP por un AR, éste lo detectaría y, siguiendo el modelo COPS outsourcing [3], preguntaría al QoS Broker qué hacer con ellos. Como éste ha recibido del SIP Proxy cómo tratar este tipo concreto de flujo (Figura 1), el QoS Broker configuraría los routers de la manera adecuada. Según se discutía en la sección 2.2 hemos visto como en Moby Dick, el SIP Proxy sólo envía una autorización al QoS Broker y la reserva de recursos se hará posteriormente; pero en nuestra red no la realiza el Terminal del usuario, sino el AR al detectar un nuevo flujo. Nótese que el dialogo del QoS Broker con los Routers es un aspecto puramente de Moby Dick, por lo tanto la mencionada limitación, realizada para simplificar el escenario de pruebas, no representa ningún impedimento para mostrar nuestra realización.

3 Conclusiones

En este artículo hemos mostrado cómo integrar servicios multimedia en una red 4G. No hemos necesitado modificar las aplicaciones que corren en los terminales de usuario (SIP UA), sólo el SIP Proxy. Hemos mostrado cómo un proveedor de red puede ofrecer, entre otros, QoS y AAA a proveedores de servicios convirtiéndose en un agregador de servicios. Además nos hemos centrado en escenarios de redes 4G y servicios de conferencia, cuando las soluciones actuales se centran en redes 3G y para servicios de provisión de contenidos. Esto se traducirá en acuerdos con proveedores de servicios para que estos servicios sean más atractivos para los usuarios frente a los ofrecidos por otros proveedores de servicios que no tengan acuerdo con el proveedor de red. Yendo aún más lejos, los propios proveedores de redes podrían ofrecerlos directamente de forma exclusiva consiguiendo todo el mercado en esa red. De esta forma los proveedores de red recuperan su papel central en le mundo de las telecomunicaciones y no se ven reducidos a ser meros transportistas de bits.

Nuestro trabajo ha sido implementado, integrado y probado en un prototipo de red 4G IPv6, Moby Dick. Se han integrado los aspectos de QoS y AAA de esta red, dejando el aspecto de movilidad para futuros trabajos. Nos atrevemos a decir que nuestra implementación es de las primeras en integrar un entorno SIP con los elementos de un proveedor de red 4G. Aunque, evidentemente, nuestro trabajo no es más que un pequeño prototipo, los resultados obtenidos son significativos. El proyecto IST Daidalos [6], en el que participan los autores de este artículo, también considera fundamental y aborda la integración de servicios SIP en redes 4G.

Agradecimientos

Este trabajo ha sido parcialmente financiado por la Comisión Europea a través del proyecto Daidalos (FP6-2002-IST-1-506997) y del proyecto SATNEX (FP6-2002-IST-1-507052).

Referencias

[1] 3GPP TS 24.228 “Signalling flows for the IP multimedia call control based on SIP and SDP (Release 5)”

[2] P. Calhoun et al. RFC 3588 “Diameter Base Protocol”, September 2003.

[3] D. Durham, Ed., “The COPS (Common Open Policy Service) Protocol” RFC 2748, January 2000.

[4] G. Camarillo ed. “Integration of Resource Management and Session Initiation Protocol (SIP)” RFC 3312.

[5] G. Camarillo and M.A. García Martín “The 3G IP Multimedia Subsystem (IMS)”, John Wiley & Sons, 2004. ISBN: 0470 87156 3

[6] IST Daidalos Project http://www.ist-daidalos.org

[7] IST Moby Dick Project. http://www.ist-mobydick.org.

[8] J. Loughney. “Authentication, Authorization, and Accounting Requirements for the Session Initiation Protocol (SIP)” RFC 3702. February 2004

[9] Pascal Kurtansky et al., “Extensions of AAA for Future IP Networks”, IEEE Wireless Communications and Networking Conference WCNC 2004. Atlanta. USA. WCNC 2004.

[10] Victor Marques et al., “An IP-based QoS Architecture for 4G operator scenarios”. IEEE Wireless Communications. June 2003

[11] NIST. IPTel Project. http://dns.antd.nist.gov/proj/iptel/

[12] S. Salsano, L. Veltri, “QoS Control by Means of COPS to Support SIP-Based Applications” IEEE Network, March/April 2002

[13] Raúl Sánchez Martín “Integración de servicios multimedia en redes 4G” Proyecto de Fin de Carrera, Universidad Carlos III de Madrid, enero 2005.

Análisis de Mecanismos para la Movilidad Transparente de Sesiones en el IMS

Luis Galindo[1], Fermín Galán[2], Miguel Gómez[2], Tomás Robles[3], Omar Walid[3], Tomás de Miguel[3], Pablo Guijarro[3]

[1]Tecnología de Redes. Telefónica Móviles España
Madrid, España
galindo_la@tsm.es

[2]Agora Systems, S. A.
Madrid, España
{fermin.galan, miguel.gomez}@agora-2000.com

[3]Departamento de Ingeniería Telemática (DIT). Universidad Politécnica de Madrid (UPM)
Madrid, España
{trobles, omar, tmiguel, guijarro}@dit.upm.es

***Abstract**. The 3GPP IMS (IP Multimedia Subsystem) is the architecture within the Packet-Switched domain of 3G core networks that provides QoS granted multimedia services (videoconference, media streaming, etc.) to subscribed users. Although IMS is part of the core network, and therefore quite independent of the access technology, cellular and wireless access networks (like UTRAN or WLAN) are the most interesting options, since users can enjoy subscribed services in any place (roaming) with session mobility capabilities. Roaming has been considered from the beginning in the design principles of IMS. However, there are some open issues that need to be solved in order to provide seamless session mobility. This paper studies several mobility-oriented scenarios and analyses different solutions for Internet environments involving SIP (Session Initiation Protocol) and MIP (Mobile IP).*

1 Introducción

Las redes móviles celulares se han convertido en un elemento habitual en nuestras actividades diarias, donde millones de personas utilizan los servicios de comunicaciones que proporcionan (voz, mensajería y, cada vez más, servicios multimedia completos) con la libertad que proporcionan la ubicuidad y movilidad inherentes a este tipo de redes. De manera particular, el éxito de la arquitectura 3GPP (*3rd Generation Partnership Project*, el organismo estandarizador de 3G para Europa) se debe a que sirve como un núcleo fundamental para el diseño de nuevas redes, que integran otras tecnologías de acceso como WLAN (*Wireless Local Area Network*) o redes fijas.

En este contexto, la movilidad llega a ser una pieza clave para proveer servicios que permitan a los usuarios estar siempre conectados. Las características básicas de *roaming* se convierten en el punto de inicio, pero la tecnología no puede quedarse ahí, ya que los usuarios demandan la continuidad de la sesión y la continuidad del servicio cuando se mueven entre redes pertenecientes a distintos operadores o cuando cambian de tecnología de acceso dentro de la red de un operador.

El IMS (*IP Multimedia Subsystem*) es un recubrimiento de señalización del dominio de paquetes del núcleo de red 3G definido como parte de la arquitectura 3GPP. Está basado en protocolos de señalización definidos por el IETF como SIP (*Session Initiation Protocol*) [1], que el propio IETF ha enriquecido en base a los requerimientos del 3GPP para soportar las características propias de las redes móviles.

En este artículo analizaremos las funcionalidades de movilidad que provee actualmente el IMS. Aunque IMS es un sistema desarrollado por el 3GPP para las redes UMTS, se caracteriza por su neutralidad tecnológica y en ese ámbito lo consideraremos a lo largo de este artículo. Por tanto, el análisis y las soluciones aquí planteadas será validas tanto para redes UMTS convencionales (con red de acceso GPRS) o integradas con WLAN, como para redes cdma2000 del 3GPP2 o, incluso, las Redes de Próxima Generación (*Next Generation Networks* o NGN), actualmente en proceso de estandarización por parte de TISPAN y que utilizan el IMS en su núcleo de red.

Hoy en día, las especificaciones de IMS no consideran la transferencia de sesión entre redes diferentes, por lo que ante un cambio de red una sesión establecida se interrumpe, siendo posteriormente necesario reestablecerla en la red visitada. En este entorno, la integración de los mecanismos de movilidad MIP (*Mobile IP*) [2] y SIP en el IMS podría proveer una buena solución para la

transferencia de sesión entre redes, ofreciendo la continuidad del servicio.

El resto del artículo se organiza del siguiente modo: la sección 2 describe brevemente la arquitectura IMS, en la sección 3 se analizan los diferentes tipos de movilidad y los escenarios que surgen de considerar la combinación de esos distintos tipos. A continuación, la sección 4 describe los mecanismos de movilidad IP dentro del contexto del IMS, poniendo el acento en la movilidad de la capa de red usando MIP y de la capa de aplicación mediante SIP. En la sección 5 proponemos diferentes posibilidades para la transferencia de sesiones, destacando la necesidad de integrar MIP y SIP para conseguir traspaso transparente. Finalmente en la sección 6 se ofrecen algunas conclusiones y líneas de trabajo futuro.

2 Arquitectura IMS

El IMS es el componente de la red 3G para la provisión de servicios multimedia basados en conmutación de paquetes con soporte de QoS (*Quality of Service*) y AAAC (*Authentication, Authorization, Accounting and Charging*). Ha sido estandarizado por el 3GPP para redes UMTS (*Universal Mobile Telecommunication System*), pero también ha sido adoptado por 3GPP2 para redes cdma2000.

La arquitectura IMS [3] (Fig. 1) especifica una capa de señalización que actúa sobre los recursos de la capa de transporte, de la que se supone banda ancha y capacidades de QoS. Aunque originalmente se desarrolló considerando GPRS (*General Packet Radio Service*) sobre UTRAN (*UMTS Terrestrial Radio Access*) en el acceso y *backbone* IP como núcleo, realmente IMS es una arquitectura neutra con respecto a la tecnología de la capa de transporte, siempre y cuando esté basada en IP. Por tanto, sería posible utilizar IMS en las redes tipo cdma2000 del 3GPP2 (como, de hecho, se recoge en sus estándares), o incluso con tecnologías de acceso WLAN o xDSL (*Digital Subscriber Line*).

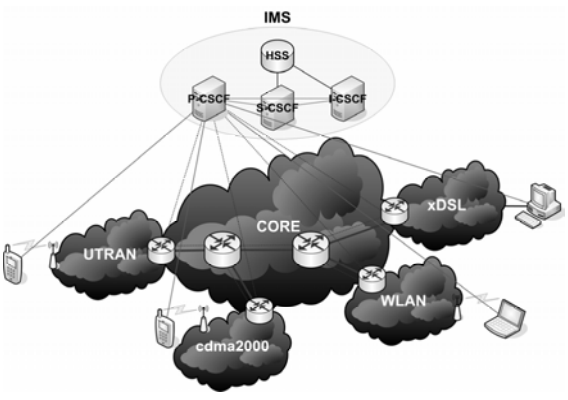


Figura 1. Arquitectura general de IMS

El subsistema IMS en sí, consta de un conjunto de nodos CSCFs (*Call Session Control Function*), que básicamente son proxies de señalización encargados de establecer, modificar y liberar las sesiones de medios con soporte de QoS garantizada y AAAC. Hay varios tipos [4]: P-CSCF (*Proxy* CSCF), que es el primer punto de contacto del terminal móvil con la red (UE en terminología IMS), el I-CSCF (*Interrogating* CSCF), que actúa como nodo frontera de la señalización entre IMSs de diferentes operadores e incluye funciones de ocultación de la topología, y el S-CSCF (*Serving* CSCF), que se encarga de registrar al usuario y del control de la sesión.

Otra entidad importante es el HSS (*Home Subscriber Server*), que se encarga de almacenar los perfiles multimedia de los usuarios, así como los datos de autenticación y autorización relativos a estos. Además, la arquitectura IMS define una interfaz SIP estándar hacia los servidores de aplicación (*Application Server* o AS), para proveer servicios avanzados.

3 Escenarios de Movilidad en Redes 3GPP

Dentro de la arquitectura definida por el 3GPP, el IMS es el subsistema implicado en el control de la movilidad de sesiones. Los agentes que pueden intervenir en la comunicación son variados, pudiendo tratarse de clientes SIP del IMS o externos a él (los cuales no soportan las extensiones SIP definidas en las especificaciones), clientes de la red de telefonía fija (PSTN, *Public Switched Telephone Network*), servidores de aplicaciones (AS), etc. En esta sección asumimos que los terminales implicados son clientes SIP conformes al IMS.

Adicionalmente, la existencia de distintos operadores gestionando su red dentro del IMS hace que se presenten distintas posibilidades, en función del operador con el que ha suscrito servicios cada usuario y de la ubicación en que se encuentra en el momento de mantener la sesión. Suponiendo que los dos terminales que desean establecer sesión no cambian de ubicación en ningún momento, se pueden definir distintos escenarios en función de las redes de los distintos operadores que estén implicadas en la sesión, como ya ha sido comentado.

Cada terminal puede conectarse al IMS desde la red del operador con el que tiene suscrito el servicio (red propia), o desde la red de otro operador con el que el primero tenga un acuerdo de itinerancia o *roaming* (red visitada). Además, los dos terminales (origen y destino de una llamada) pueden estar suscritos al mismo operador o a operadores distintos. Dependiendo de todos estos factores, caben ocho posibilidades, que se resumen en la Tabla 1.

Tabla 1. Escenarios según la localización de los terminales

SITUACIÓN DE LOS TERMINALES		DESTINO	
		Red propia	Red visitada
ORIGEN	Red propia (*Home network*)	Mismo Operador	Mismo operador
		Distintos operadores	Distintos operadores
	Red visitada (*Visited network*)	Mismo Operador	Mismo operador
		Distintos operadores	Distintos operadores

Para el análisis de la movilidad de sesiones se puede tomar como situación inicial cualquiera de las de la Tabla 1. Adicionalmente a la localización inicial del terminal que efectúa el proceso de movilidad, el destino del mismo puede ser una red de su operador o de otro operador (ver Tabla 2).

Hay que señalar que el hecho de que un terminal pase de la celda de acceso de un operador a la de otro no implica necesariamente desplazamiento físico del terminal, ya que ambas celdas pueden estar cubriendo el mismo espacio geográfico.

De la combinación de las posibilidades de la Tabla 1 y la Tabla 2 identificamos los procedimientos esenciales que van a intervenir en una transferencia de sesiones entre diferentes dominios de red y diferentes operadores. Estos procedimientos básicos serán:

- Registro de terminales, que dependerá de si es en la red propia o red visitada.
- Establecimiento de sesión, que dependerá de si es con un terminal de la misma red o de otra red y si este se encuentra en su red propia o en una red visitada.
- Transferencia de una sesión de un terminal en un dominio al mismo terminal en otro dominio diferente.
- Finalización de la sesión.
- Desregistro de los terminales.

La definición de estos procedimientos o la adaptación de los definidos por el 3GPP y su adecuada combinación permitirán realizar la transferencia de sesiones en entornos 3GPP.

La transferencia (*handover* o *handoff*) se puede clasificar de acuerdo a distintos criterios, como el número de elementos de red implicados en ella o la forma en que se realiza. La clasificación más habitual es aquella que considera el efecto sobre el servicio que tiene la transferencia, permitiendo así calificarlas como:

Tabla 2. Escenarios de movilidad de sesión

MOVILIDAD DE SESIÓN		CELDA DESTINO PERTENECIENTE A	
		Mismo operador	Distinto operador
SESIÓN MANTENIDA CON	Mismo operador	Escenario 1	Escenario 2
	Distinto operador	Escenario 3[1]	Escenario 4

[1] En este caso no es necesario un cambio de celda, simplemente el usuario decide, en un momento dado, continuar la sesión haciendo uso de los servicios contratados con otro operador sin cambiar la red de acceso.

- *Rápidas (fast handover*): cuando se minimiza el tiempo necesario para establecer los nuevos enlaces.
- *Suaves (smooth handover):* cuando no produce perdida de datos en ese tiempo.

Si se cumplen estas dos condiciones (rápida y suave) la transferencia se realiza sin que suponga un impacto perceptible para el usuario. Este caso es conocido como transferencia transparente o *seamless* y es el más completo en términos de movilidad.

Por otro lado, considerando una red de acceso GPRS, a la hora de analizar la movilidad de sesión y definir los mecanismos para conseguirla hay que tener en cuenta los siguientes puntos. En el caso de otras redes de acceso (por ejemplo, cdma2000) surgen consideraciones similares.

- El UE establece un contexto con el GGSN (*Gateway GPRS Support Node*) cuando se establece su conectividad con la red de acceso, asignándosele, entre otras cosas, una dirección IPv6.
- En general, los cambios de ubicación dentro de la red de acceso de un operador no implican cambio de GGSN y, por tanto, se mantiene el contexto establecido y, en concreto, la dirección IPv6 del UE.
- En el caso de que haya cambio de SGSN (*Serving GPRS Support Node*), se realiza una actualización (UPDATE) del contexto para indicar dicho cambio, pero sigue manteniéndose el contexto PDP (*Packet Data Protocol*).
- Por tanto, la movilidad dentro de un operador se resuelve por la propia red de acceso (IP CAN, *Connectivity Access Network*) mediante mecanismos de bajo nivel, por lo que no requiere intercambio de mensajes SIP en el nivel de aplicación.

El cambio en la dirección IP del terminal se producirá, entonces, cuando exista un cambio de GGSN, ya sea del mismo o de distinto operador, cuando alguna de las partes decida eliminar el contexto PDP establecido o cuando el usuario decida cambiar su dirección de contacto.

4 Mecanismos IP para el Soporte la Movilidad

El soporte de movilidad en redes IP puede considerarse en diferentes niveles: enlace, red (MIP [2]), transporte (mSCTP, *Mobile Stream Control Transmission Protocol* [5]) o aplicación (SIP [1]). En este punto debemos resaltar que aunque los usuarios que utilizan el IMS se benefician de los procedimientos de movilidad a nivel de enlace (por ejemplo el *handover* en GPRS en las redes 3GPP), esas facilidades de movilidad pertenecen a la red de acceso, y están fuera del ámbito de este artículo.

A continuación vamos a resumir brevemente las principales características de los elementos clave para proporcionar los servicios de movilidad en las redes IP (MIP y SIP) en el contexto del IMS, y algunas características diferenciadores de la propuesta americana (3GPP2).

4.1 Mobile IP

La principal propuesta del IETF para la movilidad a nivel de red es MIP. La Fig. 2 muestra la operación básica de este protocolo. Existen dos versiones: MIPv4 (para IPv4) [6] y MIPv6 (para IPv6) [2], que incorpora algunas optimizaciones. En ambos casos, cada MN (*Mobile Node*) mantiene una dirección IP fija denominada HoA (*Home Address*), que otros nodos (denominados CN –*Correspondent Node*) utilizarán para contactar con él, y una CoA (*Care-of-Address*) en cada red visitada, que se utiliza para lograr la conectividad en la red visitada. Existe un HA (*Home Agent*) en la red de procedencia que mantiene asociaciones entre HoA y CoA y encapsula el tráfico al MN cuando un CN quiere contactar con él (el MN se registra en el HA cada vez que su CoA cambia). MIPv4 también utiliza un FA (*Foreign Agent*), en la red visitada, que permite al MN registrar su CoA en el HA, y que también puede realizar funciones de tunelado.

El uso de MIP sin modificaciones es ineficiente Extensiones adicionales se han propuesto para resolver estos problemas, como HMIP (*Hierarchical MIP*) [7] para lograr *handovers* más eficientes minimizando la señalización fuera del dominio donde está el MN, *Binding Update* (en IPv6 viene incluido en el propio protocolo, en IPv4 se propone como extensión [8]) para resolver la ineficiencia del encaminamiento triangular, o las extensiones para el *Fast handover* (FMIP) [9], integradas con los *handover* del nivel de enlace (y que consideraremos en detalle en la sección 5.3). HMIP y FMIP han sido desarrollados para IPv6, si bien el IETF está trabajando también en soluciones para IPv4.

Como resumen, MIP proporciona servicio de movilidad transparente al nivel de transporte y de aplicación, habilitando una solución para la movilidad de terminal. Sin embargo, la movilidad de usuarios y sesiones (cuando intervienen diferentes terminales) requiere de procedimientos adicionales en el nivel de aplicación.

4.2 SIP

A nivel de aplicación, SIP permite la movilidad de los usuarios, terminales y sesiones. La movilidad personal (también conocida como movilidad de usuario) se logra asociando los servicios suscritos por los usuarios a su identidad, pudiéndose vincular ésta a una SIP URI (denominada Identidad Pública en el contexto IMS), la cual puede asociarse a cualquier dirección de contacto durante el registro SIP. El mecanismo de registro también permite la movilidad de los terminales.

La movilidad de sesiones se obtiene con las extensiones SIP REFER [10] y cabecera "Replaces" [11]. Ambos mecanismos están incluidos dentro de las extensiones normalizadas para el SIP del IMS [12]. Tanto REFER como "Replaces" son descritos con detalle en el análisis de soluciones de transferencia de sesiones que se realiza en la sección 5.2.

La movilidad basada en SIP no es transparente para el nivel de aplicación (al contrario que MIP), ya que las aplicaciones necesitan detectar los cambios de la dirección IP para disparar los procedimientos de movilidad SIP, cuando se produce un cambio de una red a otra.

4.3 3GPP2

El 3GPP2 es el organismo encargado de la estandarización de sistemas 3G en América y Asia. La arquitectura general del 3GPP2 IMS es básicamente la misma que la del 3GPP. El núcleo de red del IMS no cambia, sino que los cambios están en la red de acceso que ahora es cdma2000. Sin embargo, 3GPP2 permite dos formas de acceso (IP simple e IP móvil). Es de señalar que, a diferencia de 3GPP, que solo permite el uso de IPv6 como protocolo de red, 3GPP2 contempla el uso de IPv4.

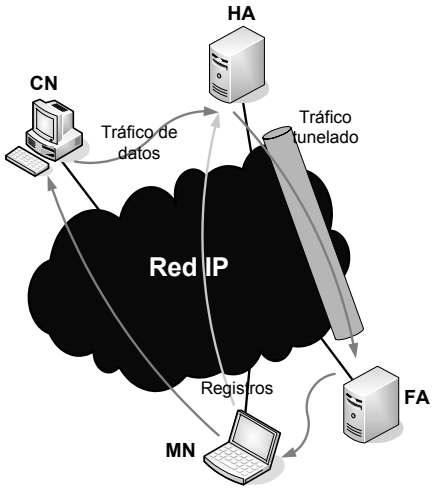


Figura 2. Operación básica MIP

La solución IP móvil está basada en MIP (descrito en la sección 4.1) y sus principales características son:

- Permite mantener la dirección IP sean cuales sean el punto y la red de acceso.
- El PDSN (*Packet Data Serving Node*, similar al GGSN en redes GPRS) ejerce ahora el papel de FA, que sirve de puente entre HA y MS (*Mobile Station*, equivalente al MN en terminología MIP).
- Se introduce la figura del HA, que centraliza el tráfico a y desde la MS.
- Para intercambiar información, FA y HA se valen de sus servidores AAA.

Las principales ventajas de la arquitectura 3GGP2 basada en MIP respecto a la del 3GPP, basada en SIP y que no contempla su utilización, son:

- Posibilidad de mantener la dirección IP cuando cambia el punto de acceso a la red, independientemente de su localización o del operador al que pertenezca. Esto permite mantener en todo momento la conectividad IP y, por tanto, las sesiones que estén activas de forma transparente en el nivel de aplicación
- Posibilidad de acceder al dominio propio desde la misma red de acceso, mediante la figura del HA. El hecho de poder cursar el tráfico de usuario en la red propia tan pronto como sea posible permite reducir los costes que supondría tener que utilizar otras redes, así como ofrecer movilidad de servicio de una forma más sencilla.

5 Transferencia de Sesiones

En esta sección se describen los distintos mecanismos que pueden ser utilizados para realizar transferencia de sesiones en el IMS.

5.1 Estado del UE en los CSCFs

Ha de tenerse en cuenta que, salvo el I-CSCF, los CSCFs son proxies SIP que mantienen ciertos elementos de estado asociados a cada UE. En concreto, cuando un UE se registra se establece un estado en el S-CSCF (dirección de registro, dirección de contacto, vector de ruta para encaminar señalización hacia el UE, criterios de filtrado, etc.). Adicionalmente, cuando un UE tiene una sesión de medios establecida, existen elementos de estado en los P-CSCF (para la recuperación ante fallos) y S-CSCF (para la recuperación ante fallos y la provisión de servicios avanzados) asociados a los UEs en la sesión de medios (Fig. 3)

Desde el punto de vista de la movilidad, el estado almacenado en los elementos de la red relacionado con el UE móvil ha de mantenerse coherente. Normalmente, esto implicará eliminar los elementos de estado asociados a la antigua ubicación, actualizándolos con la información correspondiente a la nueva ubicación.

La información de estado relacionada con el registro es relativamente fácil de mantener actualizada. El registro de un UE se almacena en un S-CSCF de la red nativa, con lo que, en el caso de que el UE cambie de ubicación (lo que implica normalmente un cambio de dirección IP y, por tanto, dirección de contacto), basta con realizar un re-registro SIP con la información actualizada.

Sin embargo, si el UE mantiene sesiones activas en el momento de realizar la transferencia, la complejidad para mantener el estado coherente es mayor, como se describe en la siguiente sección.

5.2 Mecanismos de Transferencia de Sesiones

Existen tres mecanismos que pueden ser utilizados para realizar transferencia de sesiones activas: Re-INVITE [1], REFER [10] o INVITE con cabecera "Replaces" [11].

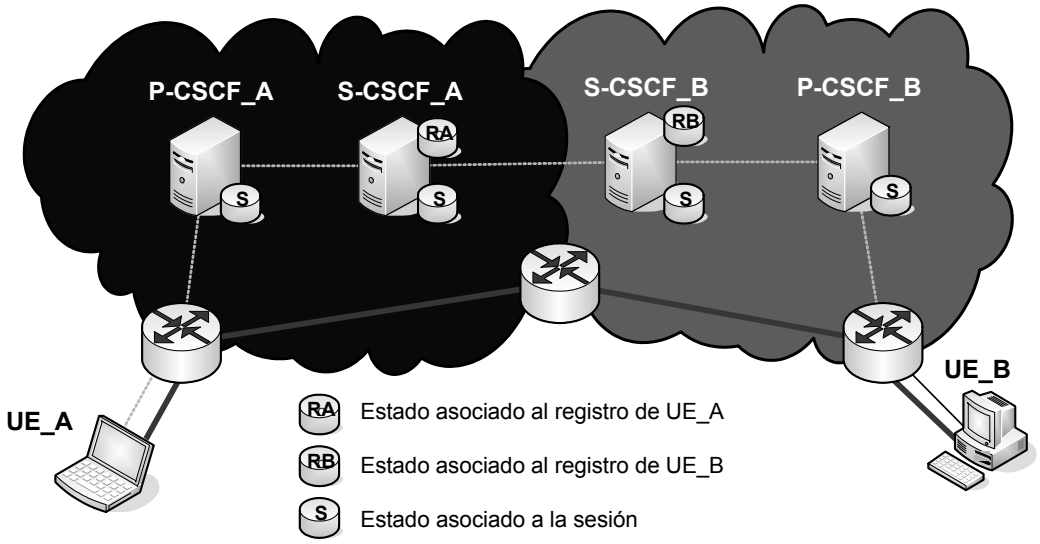


Figura 3. Estado en los CSCFs relacionado con dos UEs que establecen una sesión

El más sencillo de ellos es Re-INVITE, que consiste en el envío de un mensaje INVITE dentro del diálogo de la sesión, especificando la nueva dirección de contacto (en la cabecera "Contact"). Puesto que este nuevo INVITE sigue la misma ruta que se utilizó para iniciar la sesión (ya que va dentro de diálogo), solo sirve si la cadena de CSCFs en la nueva ubicación es la misma que en la antigua, lo cual limita bastante su utilización (posiblemente estos casos de movilidad no se den a nivel IMS y estén cubiertos por la red subyacente de transporte, GPRS en el caso del 3GPP).

En un caso general, mediante el envío de un mensaje REFER, el UE móvil puede indicar al otro UE que establezca una nueva sesión con una nueva ubicación (la ubicación a la cual el UE móvil se ha movido), especificada en la cabecera "Refer-To" (Fig. 4). El envío de REFER lleva implícito la suscripción a un paquete de eventos de estado de la transferencia, mediante el cual el UE móvil es informado de su progreso.

El UE que recibe el REFER inicia una nueva sesión y, cuando se completa el proceso de establecimiento (lo cual es notificado con un mensaje NOTIFY), el UE móvil finaliza (con BYE) la original. De esta forma, se crea el estado de la nueva sesión en los CSCFs y se elimina el viejo estado.

La alternativa a REFER es utilizar INVITE con la cabecera "Replaces" (Fig. 5). En este caso, el UE móvil inicia una nueva sesión, indicando en la cabecera SIP que la nueva sesión tiene por objeto reemplazar a otra de las sesiones que el interlocutor tiene establecidas. Cuando el procedimiento de establecimiento de la nueva sesión culmina con éxito, es el propio interlocutor quien finaliza la sesión original con un BYE. Como en el caso de REFER, el estado se mantiene coherente.

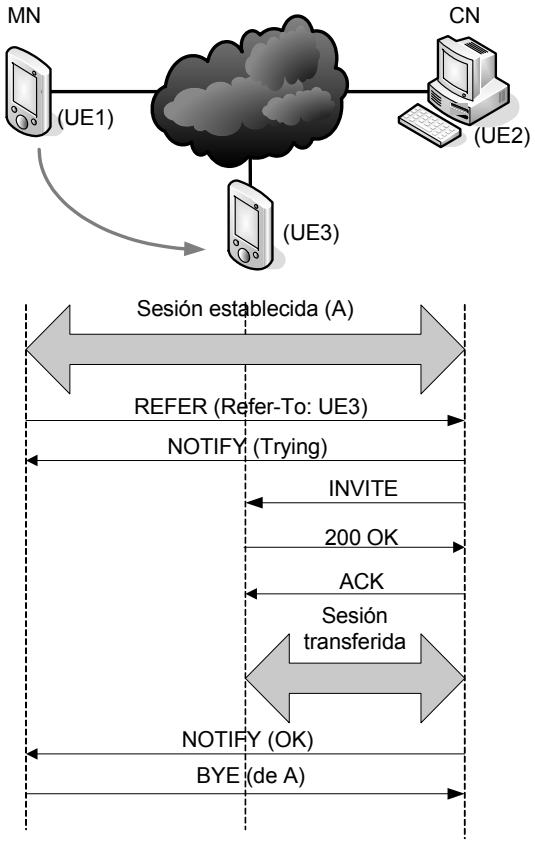


Figura 4. Transferencia de sesión con REFER

De las tres alternativas, descartamos de entrada Re-INVITE por su limitada utilidad. Entre REFER e INVITE con "Replaces", consideramos más apropiada la segunda opción, ya que resulta más eficiente (no requiere suscripciones a paquetes de eventos ni mensajes NOTIFY, y el propio mensaje de establecimiento de la nueva sesión inicia el proceso de transferencia), más segura (el uso de "Replaces" supone una autenticación implícita, mientras que REFER podría explotarse para realizar ataques de "secuestro" de sesiones) y semánticamente más correcta ("Replaces" indica que la nueva sesión sustituye a la anterior, mientras que REFER tan solo indica que se quiere establecer una nueva sesión, que no es necesariamente un reemplazo de la anterior).

5.3 Extensiones para el Soporte Transparente a la Movilidad

Los mecanismos de transferencia de sesiones descritos en la sección anterior están basados en la utilización de procedimientos de señalización SIP para lograr transferir al nuevo contexto las sesiones activas una vez se ha completado el traspaso [13]. Por tanto, estos mecanismos no permiten proporcionar movilidad transparente, ya que son reactivos (es decir, las aplicaciones han de detectar el cambio de dirección IP para poder iniciar los mecanismos de movilidad SIP al pasar de una red a otra) y relativamente pesados, ya que pertenecen al nivel de

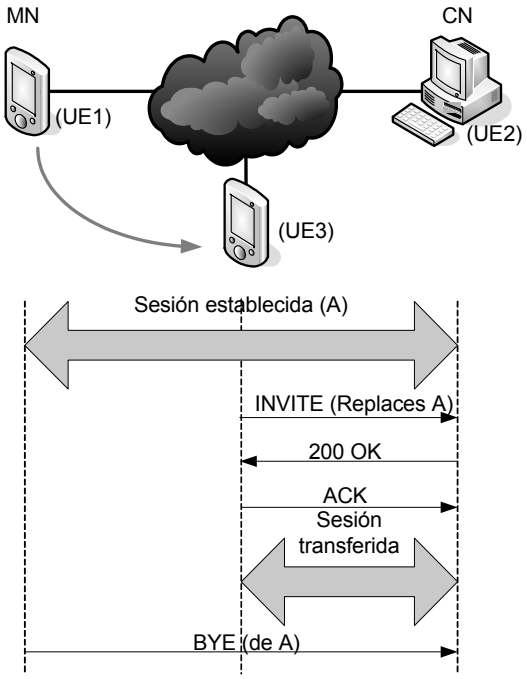


Figura 5 Transferencia de sesión con "Replaces"

aplicación. Para poder proveer servicios de movilidad que permitan el traspaso de sesiones con requisitos de tiempo real sin que se produzca un impacto perceptible en la calidad de servicio, es necesario dotar al entorno IMS de mecanismos de movilidad a nivel de red o, lo que es lo mismo, movilidad a nivel IP (MIP) [2].

Para poder beneficiarse de las ventajas de la movilidad a nivel IP, el UE móvil ha de registrarse a nivel SIP utilizando como dirección de contacto una URI basada en su HoA, ya que el hecho de detectar los cambios en la CoA e iniciar los procedimientos de movilidad SIP anteriormente descritos invalidaría por completo los mecanismos MIP. Cabría pensar entonces que, al no cambiar la dirección SIP de contacto en cada traspaso, la operación SIP de re-registro para actualizar el estado en los CSCFs descrita en la sección 5.1 dejaría de ser necesaria tras la incorporación de MIP. No obstante, dado que el mecanismo de registro SIP se emplea también para construir los vectores de encaminamiento SIP tanto en sentido terminal → red [14] como red → terminal [15], es necesario realizar este proceso cada vez que el traspaso conlleve un cambio en la cadena de CSCFs (Ej.: cuando al migrar a una nueva red, se le asignen al usuario un P-CSCF y GGSN en la red visitada), aunque el resultado del mismo sea asociar de nuevo la dirección SIP pública con la misma dirección SIP de contacto.

Igualmente, también podría pensarse que, dado que MIP permite al nodo móvil recibir y enviar tráfico independientemente de su punto de conexión a la red y de forma totalmente transparente para el nivel de aplicación, la inclusión de movilidad IP permitiría obviar los mecanismos de transferencia de sesiones descritos en la sección 5.2. No obstante, también es necesario actualizar el estado relativo a las sesiones activas almacenado en los CSCFs, ya que es de vital importancia para tareas como la tarificación, la reserva de la calidad de servicio en la red de acceso, etc. Por tanto, la inclusión de MIP en el entorno IMS no nos exime del traspaso de sesión a nivel de aplicación.

A pesar de que MIP no elimina la necesidad del re-registro y transferencia de sesiones a nivel SIP, su inclusión es imprescindible a la hora de proporcionar movilidad transparente de sesiones multimedia, ya que permite seguir recibiendo los flujos de señalización y datos en la nueva red durante el transcurso del proceso de movilidad a nivel SIP. Esto garantiza la continuidad de la sesión multimedia y la correcta terminación de los flujos de señalización relativos a la antigua red, dando una auténtica sensación de traspaso al usuario frente a la mera terminación y reestablecimiento que ofrecen los mecanismos SIP.

No obstante, MIP por sí solo no garantiza la continuidad de la sesión a nivel de tráfico multimedia, ya que los cambios de red originan un tiempo de inactividad durante el cual el UE no es capaz de enviar y recibir tráfico. Dicho periodo puede descomponerse en el retardo a nivel de enlace y el tiempo de configuración a nivel MIP.

El retardo a nivel de enlace es el tiempo de inactividad derivado del hecho de conmutar de la red origen a la red destino a nivel de enlace (ej.: cambiar de la cobertura de una red WLAN a otra usando un mismo interfaz da red). Dicho retardo es habitualmente insalvable, pero dado que se produce a bajo nivel y que existen mecanismos de optimización específicos de cada tecnología de red, suele ser lo suficientemente bajo como para resultar compatible con la continuidad de las sesiones de tráfico con requisitos de tiempo real, ya que puede ser absorbido a nivel de aplicación.

El retardo derivado de la configuración a nivel MIP es el tiempo necesario para detectar el movimiento a la nueva red, obtener la CoA pertinente, informar de la misma a la red nativa y los nodos interlocutores, etc. Dicho retardo sí resulta habitualmente incompatible con la continuidad de sesiones con requisitos de tiempo real. Por tanto, para la provisión de movilidad transparente, es necesario adoptar o soluciones basadas en *multihoming*, que eliminen por completo el periodo de inactividad del UE al permitir mantener activa la conexión simultáneamente en la nueva y la vieja red, o soluciones basadas en traspaso rápido o FMIP [9], que reduzcan el tiempo de configuración MIP hasta hacerlo compatible con la continuidad de las sesiones multimedia.

En ambos casos se produce un transitorio indeseado derivado de las condiciones de carrera entre el ACK SIP del establecimiento de la sesión en la nueva red y el mensaje MIP informando de la nueva CoA al nodo interlocutor. En el caso de que llegue primero el ACK SIP, el tráfico adopta los parámetros negociados en la nueva sesión, mientras que aún sigue siendo encaminado por la antigua red (tanto en *multihoming* como en FMIP, ya que en este último caso es tunelado de la vieja a la nueva red). En caso de que el mensaje MIP llegue primero, el tráfico comienza a fluir por la nueva red, pero aún mantiene la configuración negociada en la antigua sesión. A pesar de que dicho transitorio puede tener una duración tan breve que haga que su consideración resulte irrelevante, es posible adoptar cualquiera de las siguientes soluciones o una combinación de las mismas para solventarlo:

- Acoplar a nivel de aplicación los procesos de movilidad MIP y SIP, sincronizando así la transferencia de sesiones a ambos niveles. Esta opción tiene la ventaja de eliminar por completo el periodo transitorio no deseado, pero viola la independencia entre niveles y requiere la modificación del nivel de aplicación.
- Asumir el periodo transitorio, y proveer franquicias de tráfico en las reservas de calidad de servicio que permitan absorberlo

eficientemente (en caso de que sea un problema de establecimiento de reservas y no de capacidad absoluta del nuevo tipo de enlace) o promover el uso de aplicaciones adaptativas que detecten la reducción de ancho de banda y se comporten en consecuencia hasta que el proceso de traspaso finalice.

6 Conclusiones

En el artículo hemos revisado los diferentes mecanismos de movilidad existentes en las redes UMTS y la carencia que tienen de movilidad en el nivel de red, de forma que un cambio en la dirección IP del usuario, es decir, un cambio de red IP, produce necesariamente la pérdida de la sesión, siendo necesario que el usuario vuelva a reestablecerla.

De las tres alternativas posibles para la transferencia de sesiones activas, consideramos como más adecuado el uso de INVITE con la cabecera "Replaces" por su eficiencia, mayor seguridad y por su semántica más correcta. Pero los usuarios demandan movilidad total y transparente, lo cual no es posible con los mecanismos de transferencia de sesiones basados en SIP exclusivamente. Para ello, el uso de MIP viene a complementar la movilidad de aplicación basada en SIP.

Finalmente planteamos dos propuestas para proveer la movilidad transparente, una mediante el acoplamiento de los procesos de movilidad de MIP y SIP, y la segunda mediante el uso de aplicaciones adaptativas, capaces de detectar las reducciones de ancho de banda y actuar en consecuencia hasta que el traspaso finalice.

Si bien este artículo está centrado en aspectos teóricos, es necesaria la realización de experimentos que permitan estudiar las implicaciones de ambas propuestas, lo cual se plantea como trabajo futuro.

Agradecimientos

El trabajo descrito en este artículo se basa en los resultados de los proyectos SIUM (*Señalización Integrada para el soporte multimedia en el núcleo de red UMTS*) y EINER (*Evolución del IMS en Nuevos Entornos de Red*), una joint venture entre Telefónica Móviles de España S. A. U., Agora Systems, S. A. y el Departamento de Ingeniería Telemática (DIT) de la Universidad Politécnica de Madrid (UPM).

Referencias

[1] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., E. Schooler, "SIP: Session Initiation Protocol", IETF RFC 3261, Junio 2002.

[2] Johnson D., Perkins C., Arkko J., "Mobility Support in IPv6", IETF RFC 3775, Junio 2004

[3] 3GGP TS 23.002 v6.7.0, "Group Services and Systems Aspects; Network architecture (Release 6)", Marzo 2005.

[4] 3GPP TS 23.228 v6.9.0, "IP Multimedia Subsystem (IMS); Stage 2 (Release 6), Marzo 2005.

[5] Riegel M., Tuxen M., "Mobile SCTP", IETF draft-riegel-tuexen-mobile-sctp-04.txt, Octubre 2004.

[6] Perkins C., "IP Mobility Support for IPv4", IETF RFC 3344, Agosto 2002.

[7] Soliman H., Catelluccia C., Malki K. E., Bellier L., "Hierarchical Mobile IPv6 mobility management (HMIPv6)", IETF draft-ietf-mipshop-hmipv6-04.txt, Diciembre 2004.

[8] Perkins C., Johnson D, "Route Optimization in Mobile IP", IETF draft-ietf-mobileip-optim-11.txt, Abril 2002.

[9] Rajeev K., "Fast Handovers for Mobile IPv6", IETF draft-ietf-mipshop-fast-mipv6-03.txt, Octubre 2004.

[10] Sparks R., "The Session Initiation Protocol (SIP) Refer Method", IETF RFC 3515, Abril 2003.

[11] Mahy R., Biggs B., Dean R., "The Session Initiation Protocol (SIP) Replaces Header", IETF RFC 3891, Septiembre 2004

[12] 3GGP TS 24.229 v6.6.0 "IP Multimedia Call Control Protocol based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3 (Release 6)", Marzo 2005

[13] Schulzrinne H., Wedlund E., "Mobility Support using SIP", Proceedings of the 2nd ACM international workshop on Wireless mobile multimedia, Seattle, Washington (EE.UU.), pp. 76-82, 1999.

[14] Willis, D., Hoeneisen, B., "Session Initiation Protocol (SIP) Extension Header Field for Service Route Discovery During Registration", IETF RFC 3608, Octubre de 2002.

[15] Willis, D., Hoeneisen, B., "Session Initiation Protocol (SIP) Extension Header Field for Registering Non-Adjacent Contacts", IETF RFC 3327, Diciembre de 2002.

Integración de MPLS y DiffServ en una Arquitectura para la Provisión de QoS*

Raúl Jiménez Mateo, Cristina Paniagua Paniagua, Alfonso Gazo Cervero,
José Luis González Sánchez, Francisco J. Rodríguez Pérez
Área de Ingeniería Telemática. Departamento de Informática. Universidad de Extremadura.
Av/ Universidad s/n. 10071 Cáceres
Teléfono 927 257 253 Fax: 927 257 202
Email: raul.jimenez@zenithmedia.es,cristinap@avanade.com,{agazo,jlgs,fjrodri}@unex.es

Abstract *IP technology, even considering its success since its implantation over the Internet, suffers from limitations that make it inappropiate for a number of current telematic applications that require QoS (Quality of Service) guarantees from the network. Last years several proposals have arised trying to solve this issue. In this work it is proposed a step forward in this sense, by means of the following technologies: the adventages of the MPLS (MultiProtocol Label Switching) technology, the DiffServ (Differentiated Services) model and the use of BBs (Bandwidth Brokers). Finally, the ns-2 (Network Simulator) has been extended and used to quantify the improvements obtained by the exploitation of this new architecture, getting encouraging results.*
Keywords*— MPLS, DiffServ, QoS, Bandwidth Brokers, Ingeniería de Tráfico, Simulación de redes*

1. Introducción

En la actualidad, el protocolo IP es el dominante en la mayoría de las redes. Desde su creación, su filosofía sin conexión y con envío de tráfico de naturaleza *best effort* ha dado muy buenos resultados y, por lo tanto, contribuido a su expansión. Sin embargo, las nuevas aplicaciones que han ido surgiendo en los últimos años requieren más de lo que la actual tecnología IP puede proporcionar: altos requerimientos de ancho de banda, necesidad de transmisión con bajo retardo o sin pérdidas, etc. Para responder a estos requerimientos se han desarrollado varias formas de dotar a las redes IP de QoS (*Quality of Service*) [1]. Una de las propuestas más importantes es DiffServ (*Differentiated Services*). Otro modelo muy estudiado es la tecnología MPLS (*Multi Protocol Label Switching*), que, aunque por sí misma no proporcione QoS, es muy útil para realizar IT (*Ingeniería de Tráfico*).

Tal y como se argumentará posteriormente, una confluencia de estos dos modelos es una buena línea para mejorar las redes IP, puesto que MPLS actúa al nivel de enlace-red proporcionando un método de envío rápido por su conmutación de etiquetas y sus caminos LSP (*Label Switched Path*); y DiffServ realiza la diferenciación y priorización del tráfico necesaria para dotar a IP de QoS.

Este es el marco en el que se desarrolla nuestro trabajo, realizando un estudio que permita obtener una nueva arquitectura de red mediante la integración de los dos modelos citados (MPLS y DiffServ). Con ello se persigue, por un lado, conseguir que las redes IP permitan al usuario disponer de calidad de servicio, sin necesidad de migrar a otras tecnologías como ATM (*Asynchronuos Transfer Mode*), sin que se interrumpa el funcionamiento actual en la red y con el menor perjuicio posible para los usuarios. Además, mediante la integración de los modelos MPLS y DiffServ obtenemos una arquitectura en la que MPLS se sitúa en el nivel de red-enlace, y sirve para evitar la congestión de la red, aportando sus características de ingeniería de tráfico. Mientras, DiffServ asegura unos ciertos parámetros de calidad de servicio realizando una distinción y priorización del tráfico. Por último, la incorporación a esta arquitectura de un elemento gestor del dominio aportará ventajas como ingeniería de tráfico, optimización de recursos y control del uso de los recursos.

El documento presenta en primer lugar los trabajos relacionados, para posteriormente describir la arquitectura teórica que proponemos. Seguidamente se muestran las pruebas y los resultados obtenidos en las simulaciones realizadas para evaluar el comportamiento de la propuesta. El documento termina con una sección de conclusiones.

2. Trabajos Relacionados

La utilización de MPLS para aplicar ingeniería de tráfico promete proporcionar QoS mientras se optimizan los recursos de la red, existiendo en la actualidad un buen número de propuestas en esta línea. Sin embargo, MPLS por sí solo no puede proporcionar diferenciación de tráfico, siendo este requisito imprescindible para la provisión de garantías QoS. Por ello,

*Este trabajo ha sido financiado parcialmente por la Consejería de Infraestructuras y Desarrollo Tecnológico de la Junta de Extremadura mediante un Proyecto Regional de Investigación con referencia 2PR03A090.

puede complementarse con DiffServ para aplicar esta diferenciación. Como se expondrá posteriormente, en [2] se sugiere un mecanismo para integrar la diferenciación de servicios de forma que se traduzcan los agregados DiffServ en LSPs MPLS.

La aplicación de mecanismos de ingeniería de tráfico y de diferenciación de servicios no resulta suficiente para la provisión de garantías QoS si no se evita que la red llegue a una situación de sobreutilización de sus recursos, inevitablemente provocando congestión. Por ello, es necesario aplicar mecanismos de control de admisión, que también han sido objeto de un gran esfuerzo de investigación, encontrando propuestas tanto de aproximaciones centralizadas (como en [3]), como distribuidas (como en [4]).

Sin embargo, a pesar de las ventajas inherentes de cada una de estas tecnologías, puede apreciarse que el volumen de trabajos de investigación en torno a cada una de ellas supera ampliamente al volumen de investigación en torno al establecimiento de una sinergia que permita aprovechar lo mejor de todas ellas. Uno de los trabajos que propone la utilización conjunta de estas tecnologías es [5], aunque su propuesta no va más allá de la implementación de un *testbed* para la obtención de resultados. En [6] se presenta una arquitectura completa que, al igual que el anterior, se somete a prueba en un *testbed*. Sin embargo, la complejidad de la implementación y el requisito de implantarse completamente para ser funcional constituyen una barrera importante para su implantación en las redes actuales.

3. Arquitectura Propuesta BBArch

3.1. Clases de Servicio

En MPLS cada LSP puede estar asociado a varios FEC (*Forward Equivalence Class*), y pueden asignarse tantos flujos de información a cada FEC como sea necesario. Esto conlleva que, a efectos prácticos, pueda elegirse qué tráfico va a ser encaminado por qué LSP concreto, pudiendo implicar éste solo hecho la alteración de la QoS ofertada.

Para especificar la clase de servicio a la que pertenece cada paquete se utiliza el soporte de MPLS para DiffServ [2], donde se redefine la cabecera EXP de MPLS para la especificación de dicha clase de servicio. El campo EXP es de tres bits, por lo que cada paquete puede pertenecer a una de las $2^3 = 8$ clases posibles.

Los tipos de servicio que se proporcionan en la nueva arquitectura son los mismos que en el modelo DiffServ, ya que la integración de DiffServ con MPLS no modifica su filosofía ni su funcionamiento. De este modo se hace más sencilla la convivencia con los dominios DiffServ ya implantados.

Sin embargo mientras que en DiffServ se definen catorce tipos de servicio (EF o *Expedited Forwarding*, BE o *Best Effort* y doce tipos AF o *Assured Forwarding*), en este caso tan sólo se podrán definir ocho. El motivo de este límite puede encontrarse en [2], en el que se define E-LSP (*EXP-Inferred-PSC LSP*) como la técnica que permite que el campo EXP sea el que permita identificar la clase a la que pertenece cada paquete. Puesto que el número de tipos de servicio se ve limitado a ocho, se ha decidido implementar dos AF (1 y 2) cada una con tres niveles de descarte (1, 2 y 3), un tipo EF y otro BE.

3.2. Estructura de los nodos

La estructura básica de un nodo en la nueva arquitectura (figura 1) será la siguiente:

3.2.1. Módulo DiffServ *Pre-routing*

Clasifica los paquetes, los marca y realiza las correspondientes funciones de acondicionado del tráfico si se trata de un LER (*Label Edge Router*). Este módulo sólo se encuentra en los LERs. La figura 2 muestra el módulo de *pre-routing*, donde pueden distinguirse los distintos componentes funcionales ya explicados, así como las tablas de información y estado (perfiles y PHB) que usa. El cuadro punteado corresponde al módulo siguiente, el de MPLS a donde son reenviados todos los paquetes para la siguiente fase de proceso.

3.2.2. Módulo MPLS

Realiza las funciones de *routing* propias de MPLS, etiquetando previamente el paquete si se trata de un LER. En este bloque se traduce directamente el campo DSCP (*DiffServ CodePoint*) de IP a EXP de MPLS.

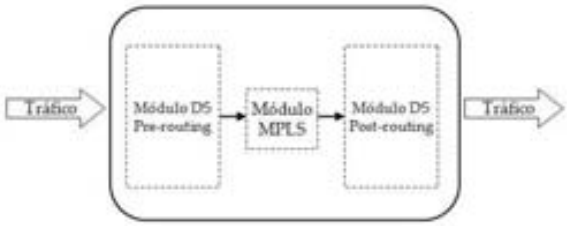


Fig. 1: Esquema de un nodo a alto nivel

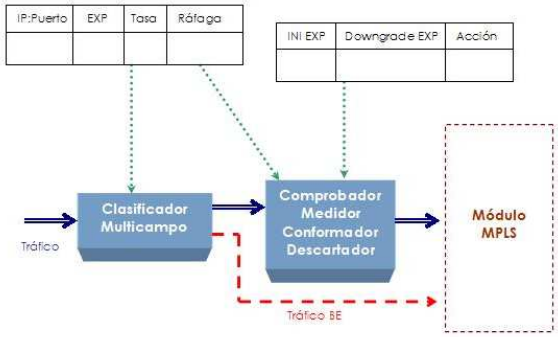


Fig. 2: Módulo de pre-routing en detalle

Además se realizan las tareas propias de *routing* de MPLS que harán que un paquete se inserte en un determinado módulo DiffServ *post-routing* u otro dependiendo de su interfaz de salida.

3.2.3. Módulo DiffServ *Post-routing*

El módulo MPLS dirige los paquetes a su interfaz de salida correspondiente. En cada interfaz se encuentra este módulo DS (figura 3), que primero realiza una clasificación por agregados de comportamiento (clasificador EXP) e inserta el paquete en la subcola adecuada, que se gestionan mediante un planificador DWRR (*Deficit Weighted Round Robin*).

Con respecto a la elección del algoritmo de planificación de colas, el principal requerimiento a la hora de elegirlo para DiffServ es que sea capaz de discriminar distintos tipos de tráfico. La política propuesta es DWRR, debido a que considera los flujos de paquetes de longitud variable, su complejidad algorítmica es baja (del orden O(1)), protege a los flujos dentro de una clase de otros flujos con mal comportamiento que puedan existir en el resto de las clases y actualmente se encuentra implementado en multitud de routers reales.

En cuanto al tipo de colas que se van a usar, todas serán de tipo FIFO, aunque se utilizarán las siguientes políticas de descarte para cada cola:

- Para la cola EF se utilizará una política simple de *Tail Drop*. Creemos que sería suficiente ya que dado que es el servicio que a priori, será el más minoritario y su prioridad es la mayor, no se esperan situaciones de congestión.
- Las colas AF serán reguladas mediante el algoritmo WRED (*Weighted Random Early Detection*) para proporcionar un mecanismo de descarte RED (*Random Early Detection*) en base a los diferentes subservicios AF con sus diferentes precedencias de descarte.
- Por último, el tráfico BE se regulará mediante el algoritmo RED.

3.3. El Bandwidth Broker en BBArch

En principio, la definición funcional del BB es similar al papel que este elemento en realiza DiffServ. Lo que cambia es fundamentalmente la interacción con los elementos del dominio. A la hora de diseñar un elemento gestor de este tipo para BBArch, una de las decisiones a tomar es si hacerlo un elemento centralizado o un sistema distribuido.

Actualmente pueden encontrarse un buen número de aproximaciones en uno y otro sentido. En la primera definición de una arquitectura para DiffServ publicada por el IETF [7] se sugiere la implementación del BB como un elemento centralizado.

Entre los problemas que generalmente se achaca a la aproximación centralizada podemos encontrar su potencial baja escalabilidad y la creación de un único punto de fallo en la red.

Fig. 3: Módulo de post-routing en detalle

Por ello, existen propuestas como [4], en la se expone un arquitectura de control de admisión completamente distribuida. Las propuestas distribuidas se basan en la dispersión de la información de estado entre todos los elementos a lo largo de la ruta de control, requiriendo por tanto protocolos más o menos complejos para conseguir esta dispersión de forma consistente. Ello acarrea necesariamente una modificación en los equipos de control instalados actualmente para soportar estos nuevos mecanismos.

Sin embargo en una propuesta centralizada la complejidad no queda diseminada por la equipación de red, reduciendo el coste del procesamiento de las peticiones de servicio y evitando generalmente la sustitución de los equipos para proceder a la implantación de la arquitectura. Debemos tener en cuenta que, para que la provisión de QoS en Internet sea una realidad, es razonable considerar una minimización la cantidad de equipación de red a sustituir.

Además, en [3] se demuestra cómo la utilización de un BB centralizado puede satisfacer los requerimientos de escalabilidad de los sistemas autónomos actuales, proponiendo a su vez la posibilidad de replicar la funcionalidad en varios equipos dispersos en la red para eliminar la existencia de un único punto de fallo, incidiendo adicionalmente en un incremento de la escalabilidad.

3.3.1. Funciones

Las funciones a realizar por el BB son una extensión de la especificación para DiffServ, donde ahora para gestionar los recursos de ancho de banda del dominio, mediante la asignación de caminos a los flujos de datos de los usuarios, debe disponer de información que le permita obtener dichas rutas a lo largo del dominio MPLS basándose en los requerimientos de calidad de servicio. En [8] se describen de forma más detallada las funciones a realizar por el BB en esta arquitectura.

3.3.2. Topología

En la definición inicial de un BB para DiffServ está contemplado que un BB debe disponer de conectividad con todos los *routers* del dominio. Esto se debe a que para proporcionar un cierto servicio en DiffServ, el BB debe configurar los *routers* para que éstos proporcionen dicho servicio. Al introducir MPLS la situación cambia, debido al establecimiento de los LSP. Por ello, no resulta necesario el control desde el BB de todos los

routers del AS (*Autonomous System*), sino tan sólo los ubicados en los bordes. Esta simplificación en la ruta de control no es, sin embargo, aplicable a la ruta de datos, en la que todos los *routers* deben implementar el módulo post-routing para la aplicación de distintas políticas de descarte a las distintas clases de tráfico.

Una vez definida la topología de la red (figura 4), podemos determinar que la comunicación será de la siguiente manera: los usuarios envían peticiones de servicio a los nodos frontera del dominio, y éstos realizan peticiones al BB. El BB da órdenes a los nodos frontera para proporcionar ciertos servicios a determinados usuarios, y los nodos lo hacen mediante la creación y/o asignación de caminos MPLS para el tráfico de éstos usuarios.

3.3.3. Petición de servicio intra-dominio

En nuestro modelo, los usuarios realizan peticiones a su LER de entrada al dominio en lugar de hacerlo al BB. Esto implica que el usuario que hace la petición debe incorporar algún mecanismo de señalización. Gracias a esta organización para la señalización, se alcanza un grado de transparencia mayor con respecto al modelo DiffServ tradicional. En el nodo frontera o LER se recibe la RAR, y se envía al BB una petición de servicio.

Una RAR (también denominada BAR o *Bandwidth Broker Allocation Request*) informa de que el usuario se dispone a realizar un envío de tráfico. Contiene la siguiente información:

- Identificador del usuario: que consta de su dirección IP, más el número de puerto fuente del nuevo flujo. Esto identifica de forma unívoca un flujo procedente de un cliente.
- Tipo de servicio que desea recibir para el tráfico que se dispone a enviar.
- Destinatario (dirección IP) del tráfico.
- *Throughput* requerido.

Cuando la RAR llega al BB, es el PDP (*Policy Decision Point*) quien se hace cargo de ella. Analiza su información, determina si es una petición correcta, e inmediatamente consulta al PMT (*Policy Management Tool*) para saber si se adecua a los datos que el BB tiene almacenados sobre ese usuario y sus SLA. Si todo es correcto y existen recursos libres para satisfacer esta demanda, se envía al usuario a través de su LER después de la apertura del LSP correspondiente, una RAA (*Resource Allocation Answer*) afirmativa, en el caso contrario se envía una RAA negativa. En el momento en que un usuario recibe una RAA afirmativa puede comenzar a enviar tráfico, pues éste será tratado con el nivel de servicio que ha especificado en la RAR (previamente contratado).

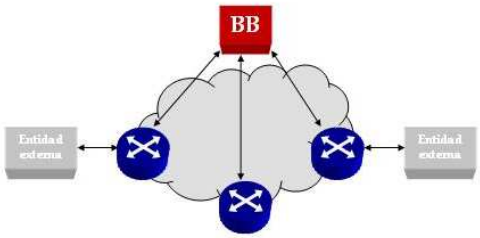


Fig. 4: Ejemplo de topología del BB

3.3.4. Petición de servicio inter-dominio

Se gestiona de igual forma que la anterior hasta el momento en el que el BB determina que el destinatario del tráfico no se encuentra en su dominio. Al descubrir esto, realiza el siguiente proceso: el BB envía un mensaje con la petición al BB par *upstream*. Éste procesa dicho mensaje y determina si el destino se encuentra en su dominio o es accesible desde él, si existen acuerdos con el BB remitente o bien si se puede satisfacer la petición con el servicio requerido, es decir, si dispone de los recursos suficientes.

Los acuerdos entre BB de dominios adyacentes se realizan previamente al proceso de admisión de SLA de clientes. Se trata de acuerdos realizados generalmente mediante un procedimiento administrativo que posteriormente son codificados en las BD de los BB, del mismo modo que los SLA de los clientes. Cuando un BB requiere que un flujo de uno de sus clientes atraviese un dominio vecino, se lo comunica a dicho dominio y éste comprueba que la petición entra dentro de la cuota reservada al dominio fuente. Por ejemplo, dos dominios adyacentes A y B pueden tener un acuerdo que consista en que el 10 % de los recursos del dominio A se reservan para el tránsito de tráfico procedente del dominio B, y viceversa.

3.3.5. Atención de peticiones y reserva de recursos

Después de que el cliente recibe una RAA afirmativa del BB puede comenzar a enviar el tráfico para el cual se envió la RAR. Previamente, el BB debe preparar el camino al nuevo tráfico. Para ello, pide al LER al que el cliente envió su RAR y que a su vez envió la petición al BB, bien que cree un LSP y envíe por él el tráfico en cuestión, que directamente lo envíe por algún LSP existente que cumpla con los requerimientos (agregación de flujos) o incluso que reubique otros flujos para dejar sitio al nuevo.

En el caso de que existan condiciones temporales en el SLA del cliente, el BB debe encargarse de que éstas se cumplan. Es decir, que si por ejemplo se ha contratado un servicio de alta calidad durante tres horas, debe controlarse el tiempo durante el que el cliente ha estado haciendo uso de este servicio, y si sobrepasa el límite de las tres horas deben denegarse los nuevos envíos de este cliente.

3.3.6. Señalización

Se requiere un protocolo de señalización que sirva para comunicar los nodos con los BBs. Puede utilizarse para este propósito algún protocolo existente, aunque la falta de estandarización en este sentido nos ha llevado a implementar uno sencillo, exclusivamente dedicado a esa tarea. Denominamos SSP (*Simple Signaling Protocol*) al protocolo de señalización que proponemos. La simplicidad del mismo resulta en una relativa simplicidad de implantación. Un protocolo de señalización complejo puede limitar la implantación y posterior expansión de la arquitectura, ya que todos los clientes deben incorporarlo.

Para el intercambio de mensajes SSP se utilizarán sesiones SSP. En este contexto se define una sesión como un diálogo entre elementos pares. Se consideran elementos pares los elementos de red que soportan el protocolo SSP y que tienen un enlace físico punto a punto entre ellos (LER y BB dentro de un dominio, BB y BB entre varios dominios). Para soportar una sesión se crea una conexión de transporte TCP durante la fase de descubrimiento. Esta conexión se mantendrá activa desde su establecimiento; y sólo se terminará en caso de que ocurra algún error fatal o se reinicie la red por algún motivo administrativo.

Aunque no es objetivo de este artículo la descripción del protocolo SSP, puede destacarse que las categorías de mensajes que se definen: *servicios*, donde se encuentran los mensajes para la comunicación de los elementos soportados por el protocolo, *topología*, destinados a los elementos pares, con los que se pueden establecer y mantener sesiones para el intercambio de mensajes SSP y *notificación.*

3.3.7. Estructura

Al igual que en DiffServ, el BB tiene dos partes diferenciadas: PDP y PMT, con las mismas funcionalidades, tal y como se refleja en la figura 5.

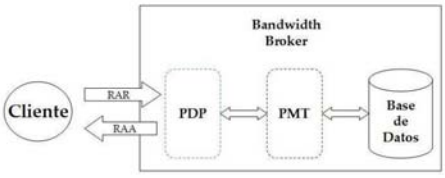


Fig. 5: Estructura del Bandwidth Broker

3.4. Especificación del SLA

En principio no existe una definición clara de qué parámetros debe contener o deben negociarse para un SLA (*Service Level Agreement*). Existen algunos estudios sobre cómo debería ser la estructura óptima de un SLA [9], pero en ningún caso se habla de un contenido estándar.

Un dato representativo de esto es que la negociación inicial, tanto en caso de SLA estáticos como dinámicos, se realiza mediante un procedimiento administrativo generalmente no automatizado. Por ello la estructura final de un SLA es finalmente una decisión que corresponde al administrador de la red e incluso puede variar dependiendo del cliente.

En cualquier caso, los parámetros de los que se debe disponer en un SLA en la arquitectura propuesta son los siguientes:

```
struct TSLA {
    int slaid;           // sla id.
    ns_addr_t uid;       // user addr.
    t_exp sid;           // service id.
    int max_throughput; // user-client
                         // max. throughput
};
```

Inicialmente supondremos que se usan SLA estáticos, es decir, que las condiciones de este contrato se negocian de forma inicial y no cambian hasta que la entidad encargada (normalmente el administrador de la red) decida realizar una nueva negociación del SLA.

3.5. Visión General de la Arquitectura

La figura 6 muestra de forma esquemática una visión global de la arquitectura, que consta de un conjunto de *routers* interconectados que forman una red y que soportan el protocolo MPLS. De éstos, los que forman la frontera del dominio se comunicarán mediante nuestro protocolo SSP con los clientes que soliciten servicios, y con el elemento que da nombre a la nueva arquitectura: el *Bandwidth Broker*. Este elemento, además de realizar las funciones propias de control de admisión y recursos (funciones que realiza en *DiffServ*), proporciona otra añadida por nuestra redefinición: La ingeniería de tráfico. La existencia de un elemento centralizado permite que, a partir de un conocimiento completo no distribuído de la asignación de recursos en el dominio, se puedan aplicar mecanismos de QoSR (*QoS Routing*) para el establecimiento de los LSP.

Por otra parte, los LER del nuevo dominio, realizan una discriminación del tráfico en clases de servicio mediante los módulos de *pre-routing* y *post-routing,* que se encuentran a la entrada y a la salida del router respectivamente. Los LSR sólo necesitan el módulo *post-routing* (además del módulo MPLS, por supuesto) lo que disminuye la complejidad en el núcleo de la red y por lo tanto, aumenta su escalabilidad, siendo coherente este punto con la arquitectura DiffServ.

4. Simulación

Partiendo de la propuesta teórica anterior se ha desarrollado, sobre el simulador *Network Simulator* (ns-2) [10], una ampliación que permita simular el comportamiento de la arquitectura propuesta. Utilizando esta ampliación, se van a presentar los resultados obtenidos de la simulación en tres pasos distintos. Se va a comenzar simulando una red IP donde no se aplique

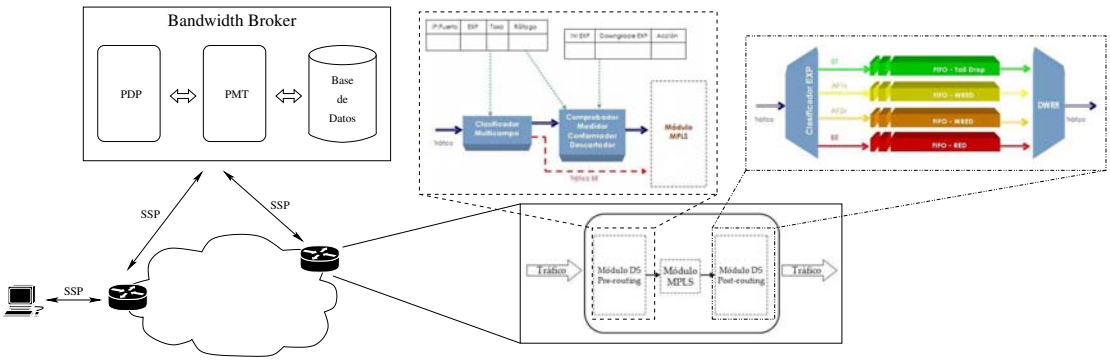


Fig. 6: Arquitectura BBArch al completo

mecanismo alguno de diferenciación de tráfico. Después se van a comparar los resultados con los obtenidos en un escenario con una topología idéntica donde se aplique la arquitectura propuesta. Por último, se examinará el comportamiento de la arquitectura cuando alguno de los usuarios incurra en un incumplimiento del SLA negociado.

4.1. Condiciones de simulación

La elección de la topología que se va a simular se ha realizado de forma que se represente una red lo suficientemente interconectada como para que exista una cierta capacidad de elección de rutas. Además, para comparar IP con BBArch, será necesario usar dos topologías muy similares, pero con la diferencia de que en la segunda estará presente el BB.

Añadir el BB supone un nuevo nodo, y un *link* que conecte dicho nodo con los LERs del dominio. La figura 7 representa la topología para la simulación de IP. Para la simulación de BBArch, la topología tan sólo añade un nuevo nodo que corresponde con el BB y que se encuentra conectado a los *router* frontera.

En los escenarios siguientes, la red va a ser atravesada por tres flujos:

- Flujo 0: Proporcionado por una fuente CBR y con unos requerimientos de QoS equivalentes a EF. La tasa contratada para EF es de 100kbps, estableciendo una tasa en ráfaga de 110kbps.

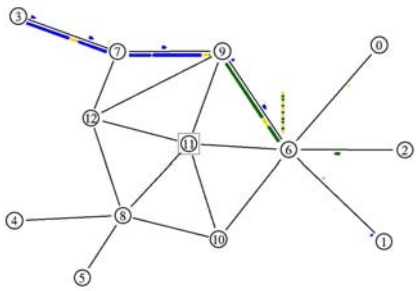


Fig. 7: Topología de la simulación

- Flujo 1: Tráfico TCP con altos requerimientos de fiabilidad (AF11). La tasa contratada para AF1 es de 500kbps, estableciendo una tasa en ráfaga de 550kbps.

- Flujo 2: Tráfico TCP, pero sin ningún tipo de pretensiones (BE).

4.2. Escenario IP

En este escenario, se simula una red en la que no se ha implantado la arquitectura BBArch y tampoco existen mecanismos específicos de DiffServ. De esta forma, se pueden comparar los resultados con los obtenidos tras su utilización. Los flujos que atraviesan la red son los descritos anteriormente y la topología utilizada se representa en la figura 7.

La gráfica de pérdidas de la figura 8 muestra claramente la ausencia de distinción entre tipos de tráfico. Además, es de suponer que dado el número de paquetes perdidos, ha habido congestiones a lo largo de la ruta, a pesar de que la red admite múltiples rutas diferentes que evitarían esta congestión. Queda de manifiesto que IP es un claro obstáculo a la hora de intentar proporcionar QoS.

Los retardos que pueden observarse en la figura 8 siguiente también eran de esperar, solapamiento, interferencias, variabilidad, todo consecuencia de la ausencia de diferenciación e ingeniería de tráfico. De hecho, las pérdidas sufridas por el supuesto EF son las más altas, algo inadmisible.

Los datos estadísticos, presentados en la tabla 1 no ofrecen la menor duda; la tasa de pérdidas del flujo 0 (supuesto EF) son mayores que las del flujo 1 (supuesto AF) y prácticamente iguales que el flujo 2 (BE).

4.3. Escenario BBArch

En este escenario, tomando la topología y tráficos del escenario anterior (con la salvedad del BB) se medirán los mismos parámetros para evaluar la mejora que supone añadir la arquitectura BBArch a una red IP.

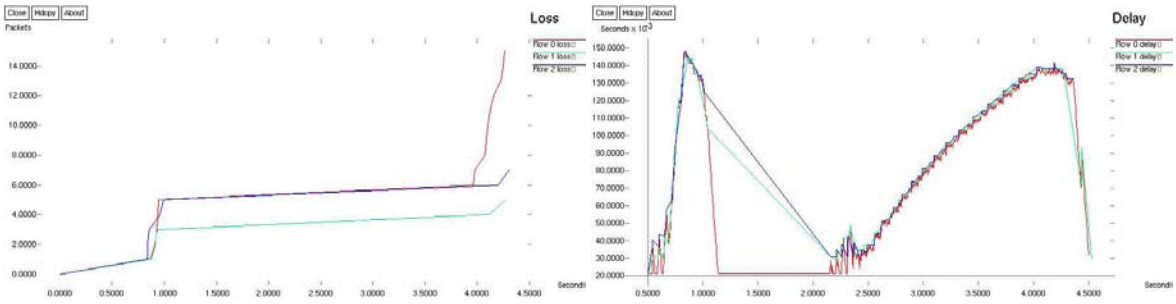


Fig. 8: Gráficas escenario IP

Tabla 1: Resultados globales de la simulación

	Escenario 0			Escenario 1			Escenario 2		
	EF	AF11	BE	EF	AF11	BE	EF	AF11	BE
Loss	1,87 %	1,28 %	1.84 %	0,00 %	0,12 %	4,89 %	67,55 %	0,12 %	7,98 %
Min Delay	21,59ms	20,99ms	21,31ms	21,49ms	26,31ms	31,63ms	27,44ms	26,31ms	32,03ms
Max Delay	147,06ms	145,01ms	148,05ms	34,23ms	82,21ms	388,93ms	37,33ms	105,33ms	307,97ms
Av Delay	84,28ms	83,00ms	84,68ms	27,86ms	54,26ms	210,28ms	32,38ms	65,82ms	170,00ms
Goodput	11,42KB/s	36,21KB/s	40,13KB/s	11,66KB/s	77,77KB/s	23,16KB/s	13,52KB/s	79,56KB/s	21,22KB/s
Goodput total	87,76KB/s			112,59KB/s			114,3KB/s		

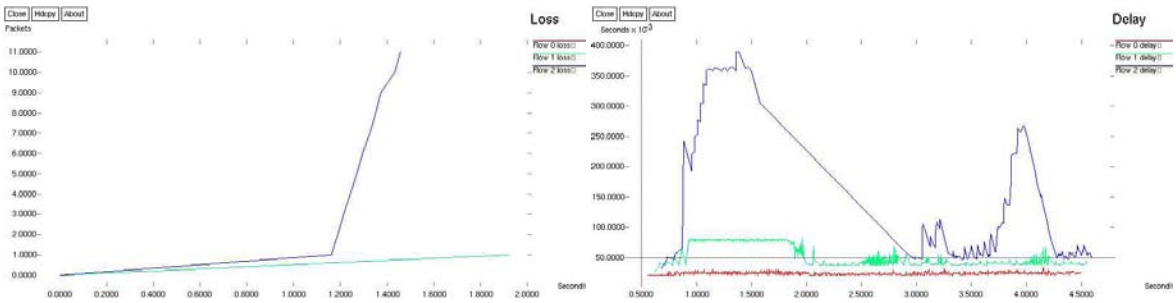


Fig. 9: Gráficas escenario BBArch

En primer lugar, en la figura 9 se observa que se han discriminado los diferentes flujos en base a su servicio contratado. Las pérdidas del flujo 0 (EF) han desaparecido y las del flujo 1 (AF) han disminuido notoriamente.

Por otro lado, el *delay* también da una idea del diferente tratamiento que han recibido los flujos en su tránsito por el dominio. Ahora, BBARCH ha conseguido que el flujo 1 no sufra el *timeout* que se producía con IP, y claramente el *delay* se ha estratificado tanto en su valor como en su variabilidad (*jitter*). El flujo 0 consigue unos excelentes parámetros de QoS, apropiados para el tráfico multimedia.

Aunque no sea algo exigible, en el flujo 1 (AF) también se observa, en la figura 9, una gran diferencia en cuanto al *delay* con respecto al escenario anterior.

Además, según puede observarse en la tabla 1, el flujo AF ha pasado de un 1.2 % de pérdidas a un despreciable 0.1 % incluso alcanzando más del doble del *throughput* que en escenario IP. Por último, puede observarse que el *goodput* se ha incrementado de manera notable, pasando de una tasa media de 87,76 KBytes/s en el escenario anterior, a 112.57 KBytes/s al añadir nuestra arquitectura.

4.4. Escenario BBARCH con incumplimiento de SLA

En este escenario se comprobará el comportamiento de la arquitectura ante la situación especial en la que los clientes incumplan el SLA negociado. Ante un incumplimiento del SLA, todo paquete EF fuera de perfil o en ráfaga, será descartado. Sin embargo el tráfico AF sólo será descartado si sobrepasa el tamaño de ráfaga. Los paquetes AF en ráfaga serán descartados (mediante EDROP) dependiendo del tipo de AF. Los resultados de la simulación se muestran en la figura 10.

En este escenario se ha incrementado la tasa a la que se envían paquetes desde la fuente CBR vinculada al servicio EF. Atendiendo a la gráfica 10, que muestra los resultados con respecto a las pérdidas, en el flujo 0 (EF) éstas han aumentado notablemente y de manera lineal. La linealidad de las pérdidas es casi con toda probabilidad producto del *metering* y el *dropping* del *TokenBucket* ya que el flujo 0 es transmitido de modo constante (CBR). Los otros dos flujos siguen dentro de

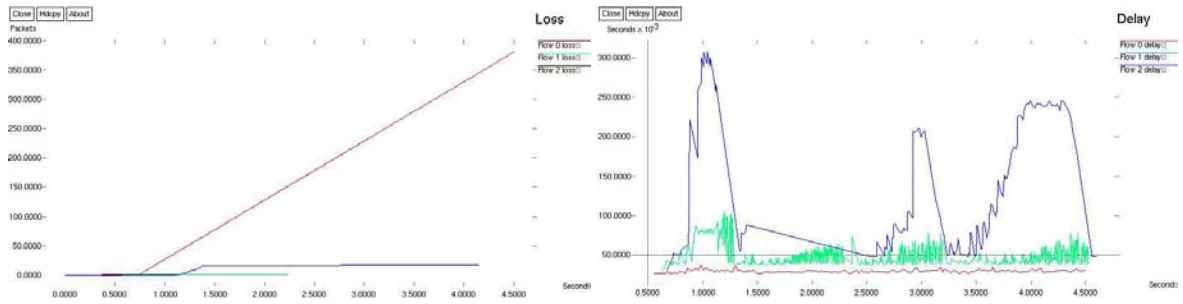


Fig. 10: Gráficas escenario BBArch con incumplimiento de SLA

lo normal. BE sigue sufriendo más pérdidas que AF.

En la figura 10, lo más sobresaliente es el aumento de los picos sufridos por el *delay* del flujo 1 debido al degradado de tráfico. A pesar de todo, no ha sufrido ningún *timeout*, debido seguramente a su alta prioridad en el planificador.

En la tabla 1 se observa un 67.55 % de pérdidas de paquetes en EF. Esto ocurre porque, cuando en EF se detectan paquetes en ráfaga, para este tipo de servicio esta situación tiene como consecuencia el descarte incondicional. A pesar de todo el delay para el tráfico EF no ha empeorado. El tráfico AF, formado por el flujo 1, ha sufrido una tasa de pérdidas muy baja. El que vuelve a sufrir un decremento en su calidad es de nuevo el flujo 2 (BE).

5. Conclusiones

El *Bandwidth Broker* es uno de los puntos fuertes de nuestra propuesta por dos razones: su importante aportación a la arquitectura, y que las soluciones de integración de MPLS y DiffServ no contaban hasta ahora con este elemento. Las funciones más importantes del BB están relacionadas con la realización del control de admisión del dominio, la gestión de recursos y la aplicación de ciertas tareas de ingeniería de tráfico.

La arquitectura BBArch auna las ventajas de las dos tecnologías sobre las que se apoya e integra: MPLS y Servicios Diferenciados. La inclusión del *Bandwidth Broker* como elemento central de la arquitectura permite un *routing* MPLS con calidad de servicio. Otros aspectos beneficiosos que añade este elemento de red son el de gestión de recursos, autentificación, control de admisión, pudiendo llegar a ofrecer servicios de tarificación y extracción de estadísticas.

Referencias

[1] Xipeng Xiao, Thomas Telkamp, Victoria Fineberg, Cheng Chen y Lionel M. Ni. A Practical Approach for Providing QoS in the Internet Backbone. *IEEE Communications Magazine*, Diciembre 2002.

[2] F. Le Faucheur, L. Wu, B. Davie, S. Davari, P. Vaananen, R. Krishnan, P. Cheval y J. Heinanen. Multi-Protocol Label Switching (MPLS) Support of Differentiated Services. IETF RFC 3270, Mayo 2002.

[3] Z. Zhang, Z. Duan y Y. T. Hou. On scalable network resource management using Bandwidth Brokers. En *IEEE Network Operations and Management Symposium*. Abril 2002.

[4] Sudeept Bhatnagar y Badri Nath. Distributed Admission Control to Support Guaranteed Services in Core-Stateless Networks. En *IEEE INFOCOM 2003*, tomo 3, páginas 1659–1669. Abril 2003.

[5] S. Avallone, M. Esposito, A. Pescape, S.P. Romano y G. Ventre. An experimental analysis of Diffserv-MPLS interoperability. En *ICT 2003. 10th International Conference on Telecommunications*, tomo 1, páginas 281–287. Febrero 2003.

[6] C. Scoglio, T. Anjali, J. C. Oliveira, I. F. Akyildiz y G. Uhl. TEAM: A Traffic Engineering Automated Manager for DiffServ-based MPLS Networks. *IEEE Communications Magazine*, 42(10):134–145, Octubre 2004.

[7] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang y W. Weiss. An Architecture for Differentiated Service. IETF RFC 2475, Diciembre 1998.

[8] Alfonso Gazo-Cervero y José Luis González-Sánchez. Incremental QoS deployment based on Network Brokers. En *HET-NETs'04. 2nd International Working Conference. Performance Modelling and Evaluation of Heterogeneous Networks*. Julio 2004.

[9] E. Bouillet, D. Mitra y K. Ramakrishnan. The Structure and Management of Service Level Agreement in Networks. *IEEE Journal on Selected Areas in Communications*, 20(4):691–699, Mayo 2002.

[10] UC Berkeley, LBL, USC/ISI y Xerox PARC. The Network Simulator - ns-2. http://www.isi.edu/nsnam/ns/.

Aplicación de Controladores Borrosos Temporales Evolutivos al Encaminamiento Adaptativo Distribuido

Manuel A. Gadeo Martos
Departamento de Ingeniería Electrónica, Telecomunicación y Automática. Universidad de Jaén.
E.U.P. Alfonso X el Sabio Nº 28.
23700 – Linares (Jaén)
Teléfono: 953 64 86 17 Fax: 953 64 65 08
E-mail: gadeo@ujaen.es

Juan R. Velasco Pérez
Departamento de Automática. Universidad de Alcalá de Henares.
Campus Universitario N-II, km. 31,5.
28871 Alcalá de Henares (Madrid).
Teléfono: 91 885 66 33 Fax: 91 885 6641
E-mail: juanra@aut.uah.es

Abstract. *In this document we carry out a comparative analysis of the application of Fuzzy Logic Controllers (hereinafter FLCs), Fuzzy Temporal Rules-Based Controllers (hereinafter FTRCs), Temporal Fuzzy Logic Controllers (hereinafter TFLCs) and Faded Temporal fuzzy Logic Controllers (hereinafter FTFLCs), to improve the adaptive distributed routing. To obtain a good knowledge bases the controllers were evolved using Genetic Algorithms.*

1 Introducción

Uno de los problemas principales en la redes de comunicaciones consiste en definir una política de encaminamiento eficiente. La política más simple hace uso de un algoritmo de encaminamiento de "camino más corto" [4], que consiste en enviar paquetes a través de los caminos compuestos por un menor número de saltos. Esta política no es óptima para todas las condiciones de tráfico en la red, debido a que se generan rutas preferidas que se congestionarán rápidamente. En está situación se formarán largas colas de paquetes, y como consecuencia un aumento en el tiempo de encaminamiento.

Muchas de las redes y protocolos asociados (por ejemplo: Arpanet, y su protocolo de encaminamiento OSPF2), emplean encaminamiento adaptativo distribuido, basado en algoritmos tradicionales como los de Dijkstra o el de Ford y Fulkerson. Estos algoritmos se ejecutan periódicamente, tomando la información sobre las condiciones del tráfico y del estado de los enlaces, que llega, procedente de los nodos adyacentes o del resto de nodo de la red. Teniendo en cuenta que el intercambio de información, de estatus de red, consume ancho de banda, el valor del periodo de actualización debe ser suficientemente grande como para no cargar en exceso la red. En estos algoritmos, las rutas que comúnmente soportan una carga fuerte son evitadas para encaminar, así como las rutas que soportan una carga débil serán seleccionadas ahora para encaminar. Lo que causa una oscilación no esperada en el proceso de encaminamiento. La naturaleza distribuida del sistema a controlar introduce dificultades adicionales al proceso de encaminamiento, achacables a que las medidas de estatus sufrirán un retraso (debido a su propagación en enlaces), y además éstas no estarán disponibles continuamente, siendo necesario realizar un muestreo a intervalos finitos [11].

En este trabajo se introduce una modificación al algoritmo de Dijkstra, utilizando un sistema de control borroso temporal difuminado, para obtener la métrica a aplicar en cada enlace, con el objetivo de mejorar el encaminamiento en redes de conmutación de paquetes no orientadas a conexión "best effort", en las que no existen garantías de calidad de servicio. Existen otros trabajos previos en los que se aplican técnicas borrosas aplicadas al encaminamiento en redes de comunicaciones. Así en [3] se propone la utilización de un algoritmo evolutivo de control temporal borroso para realizar el encaminamiento en redes de conmutación de paquetes, algoritmo que encamina los paquetes, pertenecientes a un mismo flujo, a través de las dos rutas de camino más corto, realizando una bifurcación de tráfico de manera gradual. En [1] y [9] se proponen sendas modificaciones de un algoritmo de camino más corto, a las que se les incorpora un controlador borroso, para la obtención de la métrica. En ambos casos la mencionada métrica se obtiene de forma borrosa tomando como entradas, al motor de inferencias, las características del tráfico y de los requerimiento de calidad de servicio. El objetivo que persigue dichos algoritmos [1] y [9] es la mejora del encaminamiento en redes de conmutación orientadas a conexión, con requerimientos de calidad de servicio.

El documento se organiza de la siguiente manera: en la sección 2 se realiza una aproximación a los conceptos y terminología asociada a los diferentes tipos de controladores borrosos, nombrados en el resto del documento, así como a las técnicas de aprendizaje genético propuestas. En la sección 3 se presenta el uso de FLCs (controladores borrosos) para mejorar las prestaciones de un sistema de encaminamiento adaptativo distribuido. En la sección

4 se analizan los problemas asociados con el encaminamiento utilizando FLCs, al tiempo que se incluye una justificación teórica de las mejoras introducidas por el uso de FTRCs (controladores borrosos basados en reglas temporales). En la sección 5 se muestran los problemas asociados al encaminamiento con FTRCs, al tiempo que se incluye una justificación teórica de las mejoras aportadas por la utilización de los TFLCs (controladores borrosos temporales). En la sección 6 se analizan los problemas relacionados con los encaminadores basados en TFLCs, incluyéndose una justificación teórica de las mejoras debidas al uso de los FTFLCs (controladores borrosos temporales difuminados). En la sección 7 se propone a utilización de AAGG (algoritmos genéticos) sobre FLCs ($AAGG_{FLCs}$), AAGG sobre FTRCs ($AAGG_{FTRCs}$), AAGG "Dirigidos" sobre TFLCs ($AAGGd_{TFLCs}$) y AAGG "Dirigidos" sobre FTFLCs ($AAGGd_{FTFLCs}$), a fin de optimizar el proceso de encaminamiento en las redes de comunicaciones. En la sección 8 se presenta una comparación experimental de la evaluación del comportamiento de una red de comunicaciones, simulada con distintas estrategias de encaminamiento. En la sección 9 se generaliza los resultados obtenidos en la sección 8.

2 Control borroso evolutivo

Muchas aplicaciones de los sistemas basados en conocimiento [10], necesitan manejar hechos que ocurren y varían con el tiempo. Por esta razón han surgido varios modelos para la representación y procesamiento del conocimiento temporal [2,3,4].

Los FLC son sistemas que incorporan en sus bases de conocimiento (en adelante BBC) el conocimiento humano, a través de sus reglas y las funciones de pertenencia de sus conjuntos borrosos (en adelante CCBB) [10].

Los FTRCs presentan un modelo para la representación y gestión de referencias borroso-temporales. En este modelo se define un lenguaje (con su gramática asociada) para la expresión de la información temporal de manera borrosa y se proyecta la representación de entidades temporales sobre una red de restricciones temporales borrosas [2]. Para obtener el conocimiento usado en los FTRCs, se ha empleado Algoritmos Genéticos (en adelante AAGG).

Los TFLC son un tipo particular de FLC en los que una parte del conocimiento [3,6,7,8,10] se encarga de situar, de forma borrosa en el tiempo, la aplicación de acciones sugeridas por el motor de inferencias. Para ello en la BC se incorpora la definición de los nuevos CCBB temporales, el parámetro "T" , intervalo de aplicación de las acciones sugeridas por las reglas. Al tiempo que se añade a cada regla un consecuente temporal, encargado de la ubicación temporal de las acciones sugeridas por el motor de inferencias [3].

En los FTFLCs se generaliza el modelo propuesto en los TFLCs [3], introduciendo el concepto de difuminación temporal [6,7,8], que recoge una percepción no lineal del tiempo, que dota de mayor precisión fiabilidad y certeza a las observaciones y acciones cercanas en el tiempo. Efecto que se consigue gracias a la "difuminación" de los CCBB temporales. En el motor de inferencia esta "difuminación" se traduce en una deformación de los CCBB temporales mediante unas transformaciones matemáticas, con el fin de generar los CCBB temporales transformados.

Para obtener el conocimiento utilizado en los FTFLCs y TFLCs, con un coste computacional razonable, se utilizará un método de aprendizaje genético que combina la utilización, en primer lugar, de Algoritmos Genéticos clásicos, aplicados sobre FLCs, seguido de la aplicación de AAGG sobre FTFLCs y TFLCs respectivamente . En esta última aplicación de AAGG sobre FTFLCs y sobre TFLCs, la población inicial se ha obtenido mediante clonación de la mejor BC, obtenida en el primer proceso genético (AAGG sobre FLCs). A este método de aprendizaje genético, propuesto para la obtención de las BBC, a utilizar tanto en FTFLCs como en TFLCs, se le denomina Algoritmos Genéticos "Dirigidos" sobre FTFLCs y sobre TFLCs.

3 Utilización de FLCs en encaminamiento adaptativo distribuido

La utilización de una única métrica en un sistema de encaminamiento adaptativo distribuido, es insuficiente para reflejar el estado actual de un enlace. Se trata por tanto de una limitación en la precisión de la información sobre los enlaces, que es debida al propio protocolo de encaminamiento. Así mismo la precisión de la citada información está también determinada por la amplitud del intervalo de muestreo y actualización del estado de los enlaces, de la red de comunicaciones. Para obtener mayor precisión, puede ser útil considerar dos o más métricas (variables referidas al estado de los enlaces), y asociarlas, para generar una única métrica, que describa mejor el estado de los enlaces. A tal fin, en este documento se propone la utilización de dos métricas, que serán las variables de contexto de los FLCs:

a) El valor medio del retardo sufrido por los paquetes que atraviesan un enlace, en un intervalo de observación.

b) La varianza del retardo (jitter) sufrido por los paquetes que atraviesan un enlace, en un intervalo de observación.

Valores que serán obtenidos en el intervalo de muestreo previo al instante de cálculo de la métrica (variable de operación de los FLCs). La métrica, así

obtenida, para cada enlace, tendrá un valor constante durante el próximo intervalo de muestreo. Posteriormente cada nodo ejecutará el algoritmo de Dijkstra (algoritmo de encaminamiento de camino más corto) para calcular la tabla de encaminamiento, cada T segundos (T es la amplitud de intervalo de muestreo).

4 Justificación de las mejoras introducidas por la aplicación de FTRCs al encaminamiento adaptativo distribuido.

4.1 Problemas asociados a los FLCs.

En estos sistemas de encaminamiento las rutas serán calculadas y actualizadas cada T segundos.

En un buen sistema de encaminamiento, en los que la modificación de las variables de contexto se propaga con relativa facilidad en el tiempo, será necesario ajustar adecuadamente el valor dado a las variables de operación, con el objetivo de que al finalizar el intervalo temporal de propagación, de cada acción, se consiga corregir el desajuste de las variables de contexto [6,7,8]. Por tanto, en los sistemas de encaminamiento basados en FLCs, el primer problema observado, es la imposibilidad de modificar el valor tomado por la variable de operación (para así reencaminar el tráfico) en cualquier instante.

En estos sistemas de encaminamiento basados en FLCs, las rutas que comúnmente soportan una carga fuerte (con una métrica elevada) serán evitadas para encaminar, así como las rutas que soportan una carga débil (con una métrica baja) serán seleccionadas ahora para encaminar. Hecho que causa una oscilación no esperada en el proceso de encaminamiento. Si la ruta utilizada presenta una carga elevada, el retardo medio sufrido por los paquetes crece, si se producen oscilaciones en el proceso de encaminamiento la varianza del retardo sufrido por los paquetes también crece, y por tanto la evaluación de la prestaciones que ofrece la red decrece. El segundo problema, estará asociado a la aparición de oscilaciones en el proceso de encaminamiento.

4.2 Soluciones propuestas en los FTRCs.

Para evitar el segundo de los problemas expuesto en apartado anterior, puede ser interesante tener en cuenta la información sobre el estado de los enlaces (retardo medio sufrido por los paquetes, varianza del retardo sufrido por los paquetes) no sólo en el último intervalo de muestreo, sino también en intervalos anteriores, para así poder obtener una métrica que no presente grandes oscilaciones.

Para conseguir este objetivo, se propone la utilización de un modelo de encaminamiento que utiliza un controlador borroso, basado en reglas temporales, presentado por Barro [2]. Este controlador ha sido implementado utilizando un modelo explícito de representación del conocimiento y de razonamiento. Este modelo permite explícitamente incorporar el tiempo como una variable, debido a que la evolución de las variables de contexto en el tiempo pueden ser tenidas en cuenta a la hora de realizar la inferencia.

Aplicando este controlador al proceso de encaminamiento, es posible obtener valores de la métrica, que se adapten a diferentes circunstancia, evitando la congestión de los enlaces, a la vez que grandes oscilaciones en los valores tomados por dicha métrica.

5 Justificación de las mejoras introducidas por la aplicación de TFLCs al encaminamiento adaptativo distribuido.

5.1 Problemas asociados a los FTRCs.

En estos sistemas de encaminamiento, la métrica del enlace es la variable de operación del motor de inferencia. Las variables de contexto son el retardo medio de los paquetes y la varianza del retardo, en el enlace considerado, calculadas sobre el intervalo de medida. En cada nodo se calcula la métrica de los enlaces adyacentes, cada T s., que va a ser una constante durante el intervalo de actuación de la misma. A partir de la tabla de métricas, cada nodo ejecuta el algoritmo de Dijkstra y rellena su tabla de encaminamiento, con las rutas de camino más corto, calculadas cada T segundos. El problema reside, por tanto, en la imposibilidad para modificar el valor tomado por variable de operación (reencaminar el tráfico) en cualquier instante.

5.2 Soluciones propuestas en los TFLCs.

Para salvar el problema anteriormente mencionado, es necesario que el sistema de encaminamiento pueda modificar el valor tomado por la métrica de cada enlace en cualquier instante de tiempo. Para conseguir este objetivo los TFLCs incluyen en la lista de reglas, de su base de conocimiento, una serie de reglas con idéntico antecedente a aquella que se quiere complementar, presentando en los conjuntos borrosos del consecuente, valores que provocan la modificación de los estados del sistema en el sentido adecuado (cambio en la métrica de los enlaces). Incluyéndose ahora un consecuente temporal, en el que sus CCBB tomarán valores que retrasan adecuadamente la aplicación de la acción de apoyo, permitiendo el cambio de la métrica (y como consecuencia el reencaminamiento) en cualquier instante de tiempo. Esta solución corrige en parte los desajustes, provocando un aumento en la velocidad de consecución de los objetivos marcados para las variables de contexto, mejorando así la “bondad” de la BC.

6 Justificación de las mejoras introducidas por la aplicación de FTFLCs al encaminamiento adaptativo distribuido.

6.1 Problemas asociados a los TFLCs.

El ruido externo (modificación del valor tomado por las variables de contexto , provocada por errores en el protocolo de encaminamiento), así como las reglas disparadas, durante el intervalo de aplicación, de las acciones propuesta por la variable de operación, provocarán un desplazamiento en el instante de tiempo, en el que el estado no deseado (retardo medio y varianza del retardo del enlace) que se desea corregir, se produce. Este desplazamiento dependerá especialmente de las acciones externas, generadas por un entorno ruidoso, originadas durante el intervalo de aplicación de las variable de operación.

6.2 Soluciones propuestas en los FTFLCs.

Para salvar la influencia del ruido, una buena solución puede consistir en aumentar el intervalo de actuación temporal de la acción correctora (asignación del valor de la métrica tomada por cada enlace), sin aumentar su influencia global sobre el sistema. Efecto que se consigue deformando los CCBB temporales, aumentando su base y disminuyendo su altura, siendo su área igual a la del CB original. Deformación, que para compensar la disminución en la probabilidad de ubicar correctamente la acción correctora, deberá crecer a medida que aumenta el tiempo que separa la observación del sistema y la acción programada para su control. Esta deformación que se traduce en una pérdida de certeza en la ubicación temporal de las acciones de control, conlleva un aumento en la probabilidad de ubicar adecuadamente dichas acciones.

Para poder ubicar correctamente las acciones retrasadas es necesario un número adecuado de CCBB temporales que cubra todo el "periodo", intervalo de tiempo de retardo posible ("T"). Para alcanzar este objetivo se necesita:

a) Una concentración del tiempo cercano al origen.

Teniendo en cuenta que las bases de los CCBB temporales disminuyen a medida que se acercan al origen, en esta región será necesario un mayor número de CCBB temporales, que deberán tener una menor separación entre ellos, si se quiere mantener su grado de solapamiento. Lo que implica una concentración de la región temporal próxima al origen, que se traduce en una mayor precisión en el discernimiento del tiempo así como un aumento de la certeza en la ubicación temporal de las acciones de control.

b) Una expansión del tiempo lejano al origen.

A medida que se retrasa la acción correctora aumentará la base de sus CCBB temporales asociados. Así para dos reglas que programan acciones situadas en dos instantes consecutivos alejados, los CCBB temporales estarán muy solapados, con lo que a efectos de control tendrían prácticamente la misma ubicación temporal. El objetivo marcado de programar dos acciones de control con una actuación temporal diferenciada, sólo se conseguirá si se expande el tiempo (se separa la ubicación de los instantes consecutivos). Esta expansión del tiempo deberá ser mayor al aumentar el tiempo de retardo de las reglas, para compensar el aumento del solapamiento de los CCBB temporales difuminados. Propuesta que se traduce en una pérdida de precisión en la observación del tiempo, que aumenta a medida que crece el tiempo que separa la observación del sistema y la acción programada para su control [6,7,8].

7 Aplicación al encaminamiento adaptativo distribuido del aprendizaje genético en controladores borrosos.

7.1 Estructura de las BBC.

En las BBC utilizadas en $AAGG_{FLCs}$, el conocimiento se almacena en:

a) Grupos de reglas de aplicación instantánea: conjuntos de reglas caracterizadas por presentar una única variable en el consecuente, que en todas es la misma .

b) La función de pertenencia de los CCBB de las variables de operación y contexto, caracterizada por ser funciones trapezoidales.

c) El parámetro "bondad" de la BC.

En las BBC utilizadas en $AAGG_{FTRCs}$, el conocimiento se almacena en:

a) Grupos de reglas de aplicación instantánea; grupos de reglas temporales que presentan:

a.1) Una única variable en el consecuente, que es la misma en todas las reglas (la métrica).

a.2) En el antecedente, proposiciones de la forma *"X es A <en Q de > T"*, donde *X* es una variable lingüística, *A* representa un valor lingüístico de *X*, *T* es una entidad o referencia temporal y *Q* es un cuantificador lingüístico [2].

b) La función de pertenencia de los CCBB de las variables de operación y contexto, caracterizada por ser funciones trapezoidales.

c) El parámetro "bondad" de la BC.

d) La definición de la función de pertenencia asociada al cuantificador borroso Q, caracterizada por ser trapezoidal.

e) El parámetro "Periodo", que informa sobre la longitud máxima que puede tomar la referencia temporal (T).

f) La definición de la función de pertenencia asociada a la referencia temporal (T), caracterizada por ser trapezoidal.

En las BBC utilizadas en $AAGGd_{TFLCs}$, el conocimiento se almacena en:

a) Grupos duplicados de reglas, formadas por:

1. Un conjunto de reglas no temporales.

2. Un conjunto de reglas temporales de aplicación diferida, compuesta por reglas con idéntico antecedente a alguna regla del grupo no temporal, siendo su consecuente igual o distinto a la citada regla, y a las que se añade un consecuente temporal.

b) La función de pertenencia de los CCBB de las variables de operación y contexto (forma trapezoidal), y de la variable temporal (forma triangular).

c) El parámetro "bondad" de la BC.

d) El parámetro "periodo" ("T"), intervalo máximo de actuación temporal de las reglas.

En las BBC utilizadas en $AAGGd_{FTFLCs}$, además de los elementos que componen las BBC utilizadas en $AAGGd_{TFLCs}$ el conocimiento se almacena en los parámetros de "difuminación" ("a" "b" y "c"), que modelan la variación de la precisión y certeza temporal [5].

7.2 Obtención de las BBC.

Para la implementación de los $AAGG_{FLCs}$, $AAGG_{FTRCs}$, $AAGGd_{TFLCs}$ y $AAGGd_{FTFLCs}$, se ha utilizado el enfoque de Pittsburgh, tomando como elementos de la población inicial 20 BBC, teniendo para cada caso la estructura comentada en el apartado anterior.

En la obtención de BBC aplicando $AAGG_{FLCs}$ o $AAGG_{FTRCs}$, el proceso consta de 4 fases (entre paréntesis aparece el valor dado a cada tasa):

1. Selección de BBC, proporcional a su "bondad" (0,6).

2. Entrecruzamiento de grupos de reglas (0,5) y BBC. (0,3).

3. Mutación de variables en las reglas (0,5), de CCBB (0,09) y de los puntos que los definen (0,3).

4. Sustitución por contienda de individuos viejos por nuevos, previa comparación de sus "bondades".

En la obtención de BBC aplicando $AAGGd_{TFLCs}$ o $AAGGd_{FTFLCs}$, el proceso consta de 4 fases:

1. Selección de BBC, proporcional a su "bondad" (0,6).

2. Entrecruzamiento sólo de reglas temporales, parámetro "periodo", CCBB temporales y parámetros de "difuminación" (0,3).

3. Mutación de variables en las reglas, restringida a las variables de operación y su variable temporal asociada, así como del parámetro "periodo" y los parámetros de "difuminación" (esta última mutación sólo en $AAGGd_{FTFLCs}$) (0,09).

4. Sustitución por contienda de individuos viejos por nuevos, previa comparación de sus "bondades".

8 Resultados experimentales

8.1 Descripción de la red y del tráfico ofrecido.

Para comprobar la viabilidad y conveniencia de la utilización de los controladores borrosos temporales difuminados, aplicados al encaminamiento adaptativo distribuido, se ha simulado el comportamiento de la red que se muestra el la figura 1, caracterizada por ofrecer la posibilidad de establecer caminos alternativos, para la mayor parte de los flujos de paquetes programados.

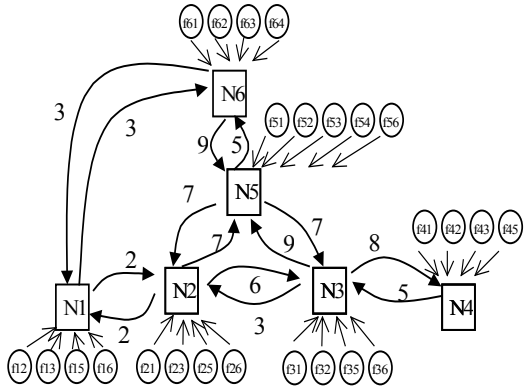


Figura 1. Modelo de red.

En este modelo se puede observar 6 nodos con 26 fuentes generadoras de paquetes (f_{ij} , "i" es el nodo origen y "j" es el nodo destino de los paquetes). Todos los enlaces presentan un ancho de banda de 10 kbps, y cada enlace se ha modelado como un sistema de cola M/M/1, siendo el número situado sobre cada enlace, la métrica inicial.

En cada fuente de paquetes, el tráfico ofrecido se modela mediante dos procesos estocásticos:

1. El Tiempo entre llegada de paquetes, que presenta distribución exponencial, con valor medio "τ", en este caso τ=0,1 s.

2. El Tiempo de servicio demandado por cada paquete, que presenta una distribución exponencial, con valor medio "s". En el ejemplo que se presenta "s" varía desde 0,02 hasta 0,09 s., y por tanto, el tamaño de paquete variará desde 200 hasta 900 bits.

8.2. Evaluación de las prestaciones de la red.

La política de encaminamiento debe orientarse a alcanzar una adecuada "calidad de servicio". En una red de conmutación de paquetes los requerimientos de "calidad de servicio" vienen dados, entre otro parámetros, por: el retardo medio sufrido por los paquetes, la varianza en el retardo (jitter), la tasa de pérdidas, el ancho de banda. En este caso sólo se tendrá en cuenta el retardo medio, el jitter y la tasa de pérdidas, para obtener así una evaluación de las prestaciones de la red. Para evaluar las prestaciones de la red se ha usado la siguiente expresión:

$$E = 0{,}7 \cdot P + 0{,}3 \cdot (1 - (0{,}8 \cdot Rn + 0{,}2 \cdot Vn))$$

E: función de evaluación. P: (paquetes llegados con éxito / total paquetes) (%). Rn: retardo medio normalizado. Vn: varianza del retardo normalizada. R: retado meio. V: varianza del retardo. TAD: retardo medio teórico, calculado para un enlace modelado con un sistema M/M/1, con una carga de tráfico de 0,8.

Para obtener una correcta evaluación de las prestaciones de la red, se simula el comportamiento de la red con:

a) En el proceso de aprendizaje:

a.1) Con cinco niveles de carga de tráfico, para todas las fuentes de paquetes. (ρ): 0.2, 0.35, 0.5, 0.65, 0.8.

a.2) Una única simulación para cada nivel de carga de tráfico.

a.3) Intervalo de simulación: 10 s.

a.4) $Rn = R / (4. TAD)$ $Vn = V / (4. TAD)$

b) En el proceso de evaluación final de las prestaciones de la red:

b.1) Con ocho niveles de carga de tráfico, para todas las fuentes de paquetes. (ρ): 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9.

b.2) Para cada nivel de carga de tráfico, se ejecutan diez simulaciones, cada una con una semilla distinta. Que son usadas para generar dos secuencias de números pseudoaleatorias, a fin de modelar el tráfico.

b.3) Intervalo de simulación: 50 s.

b.4) $Rn = R / (9. TAD)$ $Vn = V / (9. TAD)$

En el último experimento propuesto, para cada nivel de carga, la evaluación se obtiene como promedio de las diez evaluaciones obtenidas. La evaluación global se obtiene como promedio de las evaluaciones asociadas a cada nivel de carga.

8.3. Estrategias de encaminamiento utilizadas para realizar la comparación de prestaciones.

Para realizar la comparación de las prestaciones de la red, se han simulado seis estrategias de encaminamiento:

a) Encaminamiento estático de camino más corto (SPR) [5]. Los paquetes son encaminados a lo largo de rutas fijadas, para cada sesión de tráfico. Rutas que son calculadas al comienzo de la sesión aplicando el algoritmo de Dijkstra. Se trata de un método no adaptativo.

b) Encaminamiento adaptativo de camino más corto (APR) [5]. En cada nodo se ejecuta el algoritmo de Dijkstra cada T segundos, tomando como parámetros de entrada el retardo medio de los paquetes, así como la varianza del retardo, en cada enlace adyacente, obtenidos durante el intervalo de medida. Para obtener la métrica del enlace se usa una función lineal.

c) Controlador borroso evolutivo (FLC).

d) Controlador borroso basado en reglas temporales evolutivo (FTRC).

e) Controlador borroso temporal evolutivo (TFLC).

f) Controlador difuminado borroso temporal evolutivo (FTFLC) [6].

Para la obtención de los resultados experimentales:

a) Ejecutamos el proceso de aprendizaje genético para todos los controladores [6].

b) Para las seis estrategias de encaminamiento, ejecutamos el experimento propuesto, para obtener la evaluación final del comportamiento de la red de comunicaciones, con diferentes valores del intervalo de actualización T [6].

Las tablas 1, 2 y 3 muestran el resultado de una comparación representativa, de las diez realizadas, para varios valores de T.

La tabla 1 muestra el valor tomado por el parámetro bondad (evaluación de las prestaciones de la red simulada) de la BC_{FLC}, la BC_{FTRC}, la BC_{TFLC} y la BC_{FTFLC}. BBC que han sido elegidas para realizar la comparación de la evaluación final del comportamiento de la red. Dichas BBC han sido obtenidas aplicando procesos de aprendizaje genético.

Para conseguir un buen encaminamiento adaptativo utilizando FTFLCs, se deberá seleccionar, de entre las BBC obtenidas genéticamente, una BC adecuada. Esta BC deberá tomar, en sus parámetros de difuminación, valores que modelen una transformación no lineal del tiempo, así como una distorsión de los CCBB temporales, de acuerdo con el concepto de difuminación temporal [6,7].

Una comparación de la evaluación de las prestaciones de la red, para distintos sistemas de encaminamiento, es mostrada en la tabla 2, donde se puede comprobar la evaluación de las prestaciones de la red, aplicando: FLCs evolutivos, FTRCs evolutivos, TFLCs evolutivos y FTFLCs evolutivos.

Tabla 1. Bondad de las BBC utilizadas.

FLC	FTRC	TFLC	FTFLC
0,83	0,843	0,839	0,84

Tabla 2. Evaluación final de las prestaciones de la red para distintas estrategias de encaminamiento.

Carga	SPR	APR	FLC	FTRC	TFLC	FTFLC
0,2	0,889	0,489	0,964	0,970	0,973	0,972
0,3	0,694	0,313	0,812	0,861	0,876	0,877
0,4	0,593	0,247	0,711	0,738	0,739	0,745
0,5	0,525	0,220	0,606	0,642	0,636	0,649
0,6	0,471	0,175	0,534	0,567	0,560	0,572
0,7	0,426	0,195	0,485	0,511	0,501	0,517
0,8	0,387	0,137	0,441	0,471	0,459	0,479
0,9	0,352	0,169	0,396	0,431	0,423	0,437
Promedio	0,542	0,243	0,619	0,649	0,646	0,656

En la tabla 3 se muestra la mejora en la evaluación de las prestaciones de la red (en %), debida al uso de sistemas de encaminamiento FTFLCs, comparándola con la evaluación obtenida utilizando sistemas de encaminamiento SPRs, APRs, FLCs, FTRCs y TFLCs, respectivamente.

Tabla 3. Mejora introducida por la utilización de FTFLCs en comparación con otras estrategias de encaminamiento (en %).

Carga	FTFLC/SPR	FTFLC/APR	FTFLC/FLC	FTFLC/FTRC	FTFLC/TFLC
0,2	9,309	98,912	0,903	0,207	-0,094
0,3	26,391	180,375	8,044	1,860	0,145
0,4	25,716	201,846	4,762	0,941	0,892
0,5	23,613	195,389	7,144	1,157	2,083
0,6	21,492	226,599	7,066	0,898	2,116
0,7	21,334	164,354	6,455	1,191	3,072
0,8	23,798	249,375	8,635	1,721	4,244
0,9	24,393	159,005	10,342	1,483	3,480
Promedio	21,037	169,886	6,047	1,118	1,580

9. Conclusiones

Del análisis de los resultados experimentales obtenidos, se puede destacar que:

a) El uso de sistemas de encaminamiento FLC consigue una mejora en las prestaciones de la red, comparado con la utilización de sistemas de encaminamiento APR y SPR.

Mejora que es debida a que en los sistemas FLCs:

a.1) La métrica utilizada proporciona una información más completa sobre el estado de los enlaces.

a.2) En cada intervalo de actualización (T), para cada destino posible, el tráfico se puede reencaminar.

Estas dos propiedades ayudan a alcanzar el objetivo buscado, que es evitar la congestión de los enlaces.

b) El uso de sistemas de encaminamiento FTFLC consigue una mejora en las prestaciones de la red, comparado con la utilización de sistemas de encaminamiento FLC.

Mejora que es debida a que en los encaminadores FLC, durante cada intervalo de actuación de la métrica, y para cada sesión de tráfico, el encaminamiento permanece invariable. En cambio, en los encaminadores FTFLC, en cada intervalo de actuación, y para cada sesión de tráfico, pueden existir varias rutas. Con este método es posible evitar la congestión de los enlaces, en cualquier instante (no necesariamente en el instante de actualización o muestreo), pudiendo el nodo reencaminar los

paquetes a través de otro enlace, menos congestionado, sin tener que esperar.

c) El uso de sistemas de encaminamiento FTFLC consigue una mejora en las prestaciones de la red, comparado con la utilización de sistemas de encaminamiento FTRC.

Mejora que es debida a que los encaminadores FTRC, durante cada intervalo de actuación de la métrica, y para cada sesión de tráfico, el encaminamiento permanece invariable. Si bien los criterios de encaminamiento que utiliza este controlador intenta evitar la oscilación en la métrica y por tanto los fenómenos negativos asociados a ella. En cambio, en los encaminadores FTFLC, en cada intervalo de actuación, y para cada sesión de tráfico, pueden existir varias rutas. Con este método es posible evitar la congestión de los enlaces, en cualquier instante (no necesariamente en el instante de actualización o muestreo), pudiendo el nodo reencaminar los paquetes a través de otro enlace, menos congestionado, sin tener que esperar.

d) El uso de sistemas de encaminamiento FTFLC consigue una mejora en las prestaciones de la red,comparado con la utilización de sistemas de encaminamiento TFLC.

Mejora que es debida a que los sistemas FTFLCs, para resolver la incertidumbre existente sobre la información del estado de los enlaces, incorporan el concepto de difuminación en el tiempo que incluye una percepción no lineal del tiempo que da una mayor precisión, fiabilidad y certidumbre a las observaciones de los estados de los enlaces y las asignaciones de sus métricas, cercanas en el tiempo.

Referencias

[1] Aboelela E., Douligeris C., "Routing in multimetric networks using a fuzzy link cost". *2nd IEEE Simposium on computers and communications*. pp. 397-401. (1997).

[2] S. Barro, R. Marin, J. Mira, and A.R. Patón, "A model and a language for the fuzzy representation and the handling of time", *Fuzzy Sets Systems*, 61, 153-175 (1994)

[3] Carse B., Fogarty T.C., Munro A., "Artificial evolution of fuzzy rules base which represent time: a review of the temporal fuzzy classifier system", *International journal of intelligent systems*, Vol. 13, 199, pp. 905-927.

[4] D. Dubois and H. Prade, "Processing fuzzy temporal Knowledge", *IEEE Trans. Systems, Man Cybernet.*, 19, 729-744 (1989).

[5] Floyd, R.W., "Algorithm 97 (Shortest Path)". *Communications of the ACM, vol 5(6)*. 1962.

[6] Gadeo M.A., Velasco J.R., "Evolutionary strategies to achieve knowledge in faded temporal fuzzy logic controllers applied to adaptive distributed routing". *Ipmu, 2004*. Vol. 2 . pp. 789-796 .

[7] Gadeo M.A., Magdalena L., "Study of FLCs, TFLCs and FTFLCs in noisy environments", *WSES-FSFS*, 2001, pp. 5381-5386.

[8] Gadeo M.A., Magdalena L., "Ámbito de aplicación de los controladores lógico borrosos (FLC), controladores temporales borrosos (TFLC) y controladores borrosos temporales difuminados (FTFLC)", *Estylf, 2000*, pp. 309-314.

[9] Khan J. A., Alnuweiri M.A., "A fuzzy constrain-based routing algorithm for traffic engineering", *Globecom 2004,* pp. 1366-1372.

[10] Pedrycz W., "Fuzzy control and fuzzy systems", *Research Studies Press Ltd., second extended edition*, 1993.

[11] Stalling W., "Data and Computer Communications". *Macmillan*, New York, 1994.

Estudio comparativo de herramientas de evaluación de la accesibilidad Web

Vicente Luque Centeno, Carlos Delgado Kloos,
Jesús Arias Fisteus, Norberto Fernández García
Departamento de Ingeniería Telemática
Universidad Carlos III de Madrid
E-mail: vlc@it.uc3m.es

Abstract

Web Content Accessibility Guidelines 1.0 (WCAG) from W3C consist of a set of 65 checkpoints or specifications that Web pages should accomplish in order to be accessible to people with disabilities or using alternative browsers. Many of these 65 checkpoints can only be checked by a human operator, thus implying a low effective evaluation cost. However, some checkpoints can be automatically evaluated, thus spotting accessibility barriers in a very effective manner. Well known tools like HERA, WAVE, Tawdis or WebXACT evaluate Web accessibility by using a mixture of automated, manual and semiautomated evaluations. However, the automation degree of these Web evaluations is not the same for each of these tools.

This paper evaluates the WCAG coverage degree of some well known Web accessibility evaluation tools.

1. Accesibilidad

Las normas WAI (Web Accessibility Initiative) [1] del W3C han tenido un gran éxito de aceptación como guías que recomiendan mejoras en la accesibilidad de las páginas Web. La falta de esa accesibilidad en las páginas Web afecta a un elevado colectivo de la población que se encuentra con importantes barreras a la hora de navegar por sitios Web. No se trata sólo de personas con alguna discapacidad personal (dificultades en la visión, en su capacidad auditiva, en su capacidad motriz para manejar un teclado o un ratón, o en su capacidad intelectual). La accesibilidad se está demostrando como un objetivo primordial para que la publicación Web sea independiente, tanto de los dispositivos (hardware) como de navegadores o sistemas operativos (software) [16], combatiendo de esta forma la denominada "discapacidad tecnológica". De esta forma, los innumerables tipos de terminales (no sólo los basados en navegadores antiguos o alternativos, sino también los adaptados a discapacidades personales específicas de sus usuarios o los inalámbricos de reducidas dimensiones) que ofrecen capacidad de navegación en el Web, pueden llegar a ser tan funcionales como los ordenadores de mesa. La accesibilidad proporciona menores costes de mantenimiento del sitio Web y será dentro de poco (para finales de 2005 según las leyes europeas y españolas) un requisito en los sitios Web de las administraciones públicas españolas y europeas.

Sin embargo, los 65 puntos de chequeo WCAG (Web Content Accessibility Guidelines) del W3C que los documentos accesibles deben cumplir, son un conjunto muy heterogéneo de condiciones que resultan muy costosas de evaluar. La especificación de estos 65 puntos de chequeo está redactada en muchas ocasiones en un nivel de abstracción alejado de los detalles técnicos del formato HTML de las páginas Web. Muchos de esos puntos de chequeo, incluso, están abiertos a varias posibles interpretaciones subjetivas, incluyen implícitamente varias condiciones, o, simplemente, no son detectables automáticamente. Ello conlleva, como se comenta en el apartado 3 (herramientas) a que las herramientas actuales de evaluación de accesibilidad sólo puedan realizar evaluaciones parciales o incompletas.

2. Catalogación

Tradicionalmente, las herramientas de evaluación clasifican los 65 puntos de chequeo en dos grupos: los **automatizados** y los **no automatizados** (éstos últimos necesitan la comprobación de un usuario supervisor y son muy costosas de evaluar, especialmente cuando el número de páginas es elevado). Sin embargo, analizando sus detalles de implementación, se detecta que, mientras algunos criterios están claramente definidos y admiten una única interpretación (como la existencia o ausencia de determinados elementos o atributos en el marcado, según sean estos obligatorios o prohibidos), muchos otros de los puntos de chequeo automatizados están basados en criterios subjetivos y definidos de forma ambigua en el estándar original (como el hecho de que un texto sea *demasiado* largo). Conviene, por lo tanto distinguir entre:

1. Criterios **automatizables objetivos**: Son

aquellos criterios fácilmente comprobables de forma automática que están objetivamente definidos y no admiten múltiples diferentes interpretaciones. Por ejemplo, que una imagen no venga acompañada de un atributo obligatorio `alt` que encierre un texto alternativo de la imagen es algo fácilmente comprobable (aunque no lo es el hecho de comprobar que ese texto sea efectivamente una buena alternativa para la imagen).

2. Criterios **automatizables subjetivos**: Son aquellos criterios que, siendo igualmente fácilmente comprobables de forma automática, están definidos sobre condiciones *difusas* para las que ciertos grupos de personas pueden estar de acuerdo y otros grupos pueden estar en desacuerdo. Por ejemplo, la condición de que un determinado texto sea *demasiado largo* expresa una condición **difusa** donde no existe una frontera comúnmente aceptada para diferenciar lo que es demasiado largo de lo que no lo es. Esta condición, que depende de varios condicionantes externos, a la hora de ser automatizada, suele adoptar la forma de una condición no difusa (como por ejemplo que la longitud del texto no supere unos 150 caracteres), a pesar de que no esté ampliamente aceptado que un texto con 149 caracteres deba ser considerado como corto y uno con 151 deba ser considerado como largo. El W3C ha definido un conjunto de heurísticas [2] que intentan ayudar en este asunto, pero en la práctica cada herramienta tiene su propia interpretación de estas condiciones (lo cual les hace ofrecer resultados diferentes).

3. Criterios **semi-automatizables**: Son aquellos para los que el ordenador no es capaz de discernir por sí mismo un cumplimiento o un incumplimiento con un grado aceptable de confianza, razón por la cual es necesaria la intervención de un operador humano que haga ese discernimiento. Este caso se distingue del siguiente en que al menos el ordenador es capaz de **señalizar** los puntos concretos del marcado de una página que el evaluador humano debe revisar para ahorrarle el trabajo de búsqueda y focalizar así su atención sobre elementos concretos del marcado.

4. Criterios **manuales**: Son aquellos que quedan totalmente al juicio del evaluador humano y que no son en absoluto fácilmente automatizables, como el hecho de la navegación y el vocabulario sean *claros y consistentes*. Los criterios manuales y semiautomáticos, al requerir la intervención de un operador humano, son muy costosos de evaluar, especialmente cuando se trata de grandes volúmenes de documentos.

3. Comparación de herramientas de evaluación

Existen numerosas herramientas de evaluación de la accesibilidad de las páginas Web. La gran mayoría de ellas son herramientas semiautomáticas que guían al usuario a la hora de facilitarle esa evaluación. Al realizar todas ellas evaluaciones incompletas (no dan cobertura a las evaluaciones manuales), necesitan la supervisión de una persona para muchos de estos puntos de chequeo. Muchas de ellas se limitan a solicitar la atención del usuario en puntos determinados del marcado de las páginas. Nuestro estudio, se centra en herramientas completamente gratuitas y que se encuentren disponibles para ser utilizadas desde un sitio Web (quedando descartadas así las herramientas de pago o las que necesitan instalación específica en el sistema operativo del ordenador). Las herramientas seleccionadas son bien conocidas: Tawdis [14], WebXACT [15], (la sucesora de la conocida Bobby [13]), HERA [20] y WAVE [21]. Hemos tenido que excluir del estudio a Bobby [13] y Torquemada [19], porque, lamentablemente, no parece estar ya disponible para su uso.

Pese a que pudiera parecer que las tres herramientas realizan la misma función y no debieran presentar diferencias notables entre sí, las diferencias entre ellas son importantes. La ausencia actual de una normalización de los criterios, muchos de ellos subjetivos, provoca que cada una de estas herramientas tenga implementada su propia *interpretación particular* de la norma y que esas diferencias afecten, por encima de la apariencia, a la funcionalidad principal para la que fueron concebidas. Como resultado, se da el caso de que se pueden obtener resultados distintos al chequear una misma página Web dependiendo de la herramienta que se utilice. Por esa razón se recomienda hacer un **uso combinado** de varias de estas herramientas (de forma que los resultados de ellas se complementen entre sí), en lugar de centrarse en sólo una de ellas porque existen puntos cubiertos por algunas herramientas que no están cubiertos en otras. Pero, si bien esta falta de consenso resulta llamativamente grave, no lo es menos el cuantificar, como aparece detallado a continuación, el bajo nivel de **cobertura de la norma** que implementan estas herramientas respecto de lo que podrían realmente automatizar de forma no muy compleja.

3.1. Puntos de chequeo y sus condiciones

Las WCAG son un conjunto muy heterogéneo de guías. Muchas de ellas implican alguna condición que sólo el juicio humano, y no un ordenador, puede evaluar. Algunas otras implican varias condiciones como una sola. Con el fin de evaluar adecuadamente las herramientas, hemos definido las **condiciones** de los puntos de chequeo.

	Teórico	WebXACT	Taw	HERA	WAVE
Objetivamente automatizable	12	5	2	10	1
Subjetivamente automatizable	2	0	0	1	0
Semi-automatizable	32	33	20	27	15
Manual	19	27	43	27	49
Total	**65**	**65**	**65**	**65**	**65**

Cuadro 1: Comparación del grado de automatización de los puntos de chequeo del WCAG

	Teórico	WebXACT	Taw	HERA	WAVE
Objetivamente automatizable	40	25	9	23	8
Subjetivamente automatizable	11	0	0	1	0
Semi-automatizable	22	30	18	24	23
Manual	30	48	76	55	72
Total	**103**	**103**	**103**	**103**	**103**

Cuadro 2: Comparación del grado de automatización de las comprobaciones de los puntos de chequeo del WCAG

Estas condiciones de los puntos de las WCAG son un conjunto de un total de hasta 103 subcondiciones que realizan evaluaciones más detalladas y concretas que los 65 puntos de chequeo de los WCAG originales. Un punto de chequeo puede implicar una o varias de estas condiciones. Por ejemplo, el punto de chequeo 4.1 (identificar claramente los cambios de idioma) puede implicar las condiciones 4.1a (usar el atributo `xml:lang` en aquellos elementos con idioma diferente al del contexto) y 4.1b (usar el atributo `hreflang` en los enlaces que apuntan a documentos en un idioma diferente).

El cuadro 1 indica cómo las distintas herramientas de evaluación clasifican los 65 puntos de chequeo de los WCAG en las categorías anteriormente mencionadas. Dado que tanto las reglas semi-automatizadas como las reglas manuales implican un elevado coste de evaluación, nuestro interés se centra principalmente sólo en las reglas automatizadas. De un total de 12 puntos de chequeo objetivamente automatizables, HERA detecta 10 posibles, WebXACT sólo detecta 5, 2 son detectables por Taw y apenas 1 por WAVE. De ello se deduce que al menos 12 - 10 = 2 puntos objetivamente automatizables tienen una automatización mejorable en la mejor de las herramientas. Las reglas subjetivamente automatizables son tratadas normalmente en estas herramientas como semi-automatizables debido a que carecen de capacidad para especificar **preferencias** acerca de cómo evaluar esas condiciones respecto a los criterios de un usuario, con lo que acaban requiriendo la intervención humana en cada una de sus apariciones o acaban teniendo diferentes interpretaciones. Lamentablemente ninguna de estas herramientas presenta la posibilidad de establecer esas preferencias.

Los datos del cuadro 1 no reflejan sin embargo el detalle de la cuestión, pues hace referencia a la *gruesa* granularidad de los 65 puntos de chequeo originales. En el cuadro 1 se ha considerado automatizado un punto de chequeo de los 65 posibles sólo si *todas y cada una* de las comprobaciones que lo forman se encuentra completamente automatizada, lo cual es raramente posible. Lo habitual es que dado un punto de chequeo, algunas de sus condiciones sí sean automatizables mientras otras no lo sean. De ahí los números tan poco esperanzadores del cuadro 1. Pero herramientas como Taw y HERA no se limitan a detectar apenas tan pocas comprobaciones como aparecen en el cuadro 1, sino que realizan muchas otras que, al menos parcialmente forman parte de otros puntos de chequeo. Para dar una visión más detallada de la realidad de estas herramientas, esto es, con una granularidad más fina, hemos representado en el cuadro 2 los resultados aplicados a las 103 comprobaciones concretas.

En el cuadro 2 se repiten los resultados de la evaluación anterior, esta vez teniendo en cuenta las 103 condiciones encontradas en lugar de los 65 generales. Aunque a primera vista puede detectarse cómo el número de condiciones automatizadas por las herramientas es considerable (entre 8 y 25), y que la gran mayoría de las páginas diseñadas sin criterios de accesibilidad suelen incumplir al menos alguna de esas condiciones automatizadas, lo cierto es que el grado de cobertura de lo que en realidad podrían llegar a automatizar estas herramientas deja bastante que desear, pues de las 103 condiciones, 40 son fácilmente automatizables desde un punto de vista teórico (no necesitan de un algoritmo excesivamente complejo), cifra muy superior a lo que realmente ofrecen estas herramientas. Puede verse cómo el número de comprobaciones manuales, que desde un punto de vista teórico podría estar reducido a un número de 30, es sensiblemente más elevado en las herramientas (desde las 48 de WebXACT y Bobby hasta los 76 de Taw).

WCAG #	Regla	WebXACT,Taw,HERA	WAVE
1.1e	Sección alternativa para no-marcos	Auto Obj.	Manual
1.5	Texto alternativo para cada `area`	Auto Obj.	Manual
7.2a	No usar etiqueta `blink`	Auto Obj.	Semi
7.3a	No usar etiqueta `marquee`	Auto Obj.	Semi
12.4a	No usar etiquetas `label` sin asociar	Manual	Auto Obj.
13.2b	Obligación de usar título del documento	Auto Obj.	Manual

Cuadro 3: Reglas con comportamientos diferentes según la herramienta (WAVE frente al resto)

WCAG #	Regla	WebXACT,HERA,WAVE	TAW
7.4	No usar auto-refresco	Semi	Manual
7.5	No usar auto-redirección	Semi	Manual
12.4b	Todo campo de formulario debe tener descriptor	Auto Obj.	Manual

Cuadro 4: Reglas con comportamientos diferentes según la herramienta (Taw frente al resto)

3.2. Puntos de chequeo con cubrimientos comunes

De los cuadros 1 y 2, se determina que, a pesar de que estas herramientas pueden tener comportamientos comunes (o similares) para ciertos puntos de chequeo, para otros, existen notables diferencias. La siguiente lista muestra las condiciones de puntos de chequeo para las que todas las herramientas evaluadas presentan un comportamiento similar. Como era de esperar, la muy bien conocida regla 1.1a (proporcionar un atributo `alt` en todas las imágenes), figura dentro de ellas. Sin embargo este conjunto llama la atención que este conjunto esté limitado a estas dos condiciones.

- Regla 1.1a: Todas las imágenes tienen que tener texto alternativo (atributo `alt`)
- Regla 1.1b: Todos los botones imagen tienen que tener texto alternativo (atributo `alt`)

3.3. Puntos de chequeo con cubrimientos distintos

Cada herramienta realiza o deja de realizar algunas comprobaciones que son genuinas de esa herramienta. El cuadro 3 muestra algunos comportamientos particulares en los que WAVE se distingue de las demás herramientas. Como puede verse, la mayoría de ellas se trata de comprobaciones que WAVE no realiza y que delega en el usuario, mientras el resto de herramientas sí las comprueba. Nótese que, aunque WAVE no ha implementado ciertas comprobaciones realmente sencillas (como la de que el documento deba tener título o que los elementos `area` deben tener texto alternativo), sin embargo es la única herramienta que implementa la regla 12.4a (un poco más compleja) que persigue el uso de elementos `label` sin asociar a campos de formularios.

De la misma forma, el cuadro 4 muestra algunos comportamientos particulares en los que Taw se distingue, a su vez, de las demás herramientas. En esta ocasión, se trata en todos los casos de comprobaciones que Taw no realiza y que delega en el usuario, mientras el resto de herramientas sí las comprueba.

Las faltas de comprobaciones de Taw y WAVE no se limitan a las expresadas en los cuadros 3 y 4. En el cuadro 5 pueden verse más fallos, esta vez comunes a ambas herramientas Taw y WAVE, pero que sí comprueben las herramientas WebXACT y HERA.

Como ha podido desprenderse de los cuadros anteriores, las herramientas Taw y WAVE presentan varias deficiencias, mientras que las herramientas WebXACT y HERA se encuentran en una categoría superior y realizan un mayor cubrimiento de las reglas de accesibilidad. El cuadro 6 muestra las diferencias entre WebXACT y el resto de herramientas analizadas. En este caso no se trata de deficiencias, sino de comprobaciones que sólo WebXACT realiza y que las otras herramientas dejan sin implementar delegando esa responsabilidad al usuario. Entre esas comprobaciones se analiza el hecho de que los eventos dependientes del teclado deben aparecer emparejados junto con los eventos dependientes del ratón (reglas 9.2a, 9.2b y 9.2c) y el hecho de que los textos de los enlaces no deben aparecer repetidos, siendo así ambiguos (regla 13.1a). La herramienta WebXACT también reclama comprobar el adecuado uso de las cabeceras (reglas 3.5). Sin embargo, en la elaboración de estos tests se ha comprobado que esas comprobaciones sólo funcionan en un conjunto muy concreto de condiciones y que no siempre se verifican adecuadamente como cabría esperar.

WCAG #	Regla	WebXACT,HERA	Taw,WAVE
1.1c	Texto alternativo debe ser corto	Semi	Manual
4.3	El idioma del documento debe estar declarado	Auto Obj.	Manual
10.4a	Los campos no ocultos deben tener valor por defecto	Auto Obj.	Manual
10.4b	Los `textarea` deben tener valor por defecto	Auto Obj.	Manual
10.5	Debe haber texto imprimible entre enlaces	Auto Obj.	Manual
12.1	Los marcos deben tener título	Auto Obj.	Manual

Cuadro 5: Reglas con comportamientos diferentes según la herramienta (WebXACT y HERA frente a Taw y WAVE)

WCAG #	Regla	WebXACT	HERA,Taw,WAVE
3.5a	Cabeceras h2 deben ir precedidas de `h1` ó `h2`	Pseudo	Manual
3.5b	Cabeceras h3 deben ir precedidas de `h2` ó `h3`	Pseudo	Manual
3.5c	Cabeceras h4 deben ir precedidas de `h3` ó `h4`	Pseudo	Manual
3.5d	Cabeceras h5 deben ir precedidas de `h4` ó `h5`	Pseudo	Manual
3.5e	Cabeceras h6 deben ir precedidas de `h5` ó `h6`	Pseudo	Manual
9.2a	El evento `onmousedown` debe emparejar con `onkeydown`	Auto Obj.	Manual
9.2b	El evento `onmouseup` debe emparejar con `onkeyup`	Auto Obj.	Manual
9.2c	El evento `onclick` debe emparejar con `onkeypress`	Auto Obj.	Manual
13.1a	No debe haber enlaces con textos ambiguos	Auto Obj.	Manual

Cuadro 6: Reglas con comportamientos diferentes según la herramienta (WebXACT frente al resto)

Finalmente, el cuadro 7 incluye las comprobaciones que, genuinamente, solo la herramienta HERA comprueba y que las demás herramientas no cubren. Estas comprobaciones hacen referencia al uso de eventos dependientes del ratón que no son combinados junto con eventos independientes del dispositivo y al hecho de que el JavaScript no debe obstruir la navegación (de tal forma que si se encuentra deshabilitado el uso de ese lenguaje, los elementos de la página sigan siendo funcionales).

Regla 3.2: Validar documentos con una gramática pública

La validez de los documentos Web frente a una gramática comúnmente aceptada, como es XHTML [7] es un punto clave dentro de la accesibilidad. Sin embargo, ambos conceptos son lo suficientemente independientes entre sí como para que las herramientas de evaluación de accesibilidad no realicen actualmente [17] procesos de validación frente a ningún DTD o XML Schema conocido. Solamente HERA sí realiza una validación del marcado XHTML y de las hojas de estilo CSS y para ello usa los validadores del W3C [11] y [12]. Sin embargo, debemos decir que esta regla 3.2 de las WCAG es, sin duda alguna, la que implícitamente garantiza más accesibilidad que ninguna otra regla. Ello es así debido a que algunas reglas de accesibilidad están ya recogidas en las restricciones de la gramática de XHTML. De hecho, hemos encontrado que, dependiendo de las diferentes versiones de XHTML, la accesibilidad puede ser más difícil o más fácil de conseguir. Por ejemplo, XHTML 2.0 [10] proporciona más funcionalidades de accesibilidad que XHTML Basic 1.0 [8], que a su vez proporciona mejor accesibilidad que XHTML 1.1 [9], quien a su vez mejora la accesibilidad de anteriores versiones de XHTML. [7]

Entre los ejemplos de puntos de chequeo de accesibilidad que se consiguen al validar respecto a alguna versión de XHTML figuran:

- Regla 3.3: Usar hojas de estilo para controlar la maquetación y la presentación
- Regla 3.6: Usar adecuadamente el marcado de las listas y sus elementos
- Regla 11.2: Evitar el uso de funcionalidades anquilosadas de las tecnologías del W3C

3.4. Puntos de chequeo sin cubrimiento

Los cuadros de la sección 3.3 proporcionan un resumen de las diferencias más importantes en la cobertura que de las normas WCAG tienen las herramientas de evaluación seleccionadas. Por otro lado, a continuación se muestra la lista de todas las reglas que, a pesar de ser fácilmente automatizables, han sido ignoradas por todas las herramientas evaluadas en el estudio. Entre esas comprobaciones *faltantes* en las herramientas, destacan:

- Regla 4.2a: Uso no ambiguo de acrónimos y abreviaturas. Debería comprobarse que no existe más de una definición para cada texto, bien por abreviatura o por acrónimo, pues ello produciría ambigüedad en la interpretación del término. Por ejemplo si el término *UN* se define en un mismo documento tanto como *Unified Notation* como *United*

Nations existirá ambigüedad a la hora de interpretar el término.

- Regla 9.4a: Usar adecuadamente los atributos `tabindex` que regulan el adecuado orden de tabulación. Los atributos `tabindex`, si se usan, deben tener valores numéricos enteros positivos, únicos y consecutivos. Ninguna herramienta comprueba eso.

- Regla 9.5a: Usar adecuadamente los atributos `accesskey` que regulan los atajos de teclado. Los atributos `accesskey`, si se usan, deben tener un único carácter y éste debe ser único. Ninguna herramienta comprueba eso.

- Regla 12.3a: Agrupar los campos de formulario mediante el elemento `fieldset`. Ninguna herramienta requiere este elemento, probablemente porque no se le considera un elemento obligatorio, pero que lo será en las futuras versiones de XHTML.

- Regla 12.3b: Agrupar las opciones de las cortinillas de selección (elementos `select`) mediante el elemento `optgroup`. Ninguna herramienta requiere este elemento.

- Regla 12.3c: No usar párrafos demasiado largos. Ninguna regla comprueba esto debido a la subjetividad. No obstante, sería deseable incluir esa subjetividad en unas preferencias parametrizables.

A continuación se proporcionan soluciones de implementación para esas reglas que no están cubiertas por ninguna de las herramientas estudiadas. Para esta implementación basta un motor XQuery [5], un transformador HTML a XHTML como por ejemplo Tidy [18] y unas funciones subjetivas (representadas en itálicas) parametrizables mediante preferencias.

Regla 4.2a: Usar abreviaturas y acrónimos con propiedad

Las abreviaturas y acrónimos no deben ser usados de forma inconsistente. Al contrario, deben tener definiciones claras y únicas. La figura 1 detecta las abreviaturas y acrónimos que proporcionan más de una definición para un mismo texto. Esta regla, la 4.2a, no garantiza sin embargo toda la regla 4.2, puesto que no garantiza que todas las posibles abreviaturas y acrónimos sean marcados adecuadamente. Sólo comprueba que las que lo están, lo estén con propiedad.

```
(//abbr | //acronym)[let $a:=self::node() return
count((//abbr | //acronym)[text() = $a/text()])
!= 1]
```

Fig 1: Expresión XQuery para detectar las abreviaturas y acrónimos que rompen la regla WCAG 4.2a

Regla 9.4a: Especificar un adecuado orden de tabulación

Siempre que se especifique explícitamente un orden de tabulación diferente al de por defecto, los atributos `tabindex` deben ser usados de forma consistente. La figura 2 direcciona todos los elementos que tienen un atributo `tabindex` inadecuado, es decir, que no es un número adecuado o que esté compartido por varios elementos (debe ser único).

```
//*[@tabindex][let $n:=self::node()/@tabindex re-
turn not(isnumber($n)) or count(//*[@tabindex=$n])
!= 1 or number($n)<1 or number($n) >
count(//*[@tabindex])]
```

Fig 2: Expresión XQuery para detectar los elementos que rompen la regla WCAG 9.4a

Regla 9.5a: Proporcionar adecuados atajos de teclado

Cuando que proporcionan atajos de teclado mediante el atributo `accesskey`, se debe hacer de forma consistente. La figura 3 direcciona todos los elementos que tienen un atributo `accesskey` que no sea un carácter o que esté compartido por varios elementos (debe ser único dentro del documento). No se trata de exigir aquí que los enlaces importantes vayan acompañados de un atajo de teclado, sino que los elementos que dispongan de ese atajo de teclado lo tengan definido correctamente.

```
//*[@accesskey][let $c:=self::node()/@accesskey re-
turn not(ischar($c)) or count(//*[@accesskey=$c])
!= 1]
```

Fig 3: Expresión XQuery para detectar los elementos que rompen la regla WCAG 9.5a

Reglas 12.3: Dividir los bloques de información en grupos manejables

Los elementos `fieldset` son altamente recomendados para formularios que tienen varios campos editables en su interior. Pueden ser usados para agrupar varios de ellos en grupos que semánticamente tengan algún tipo de relación, proporcionando de esta forma grupos más manejables de campos editables. Aunque esta regla puede ser considerada como subjetiva (la función *toomany_inputs* debe ser definida con criterios subjetivos), tanto WebXACT como Taw tratan a esta regla como semi-automatizada, delegando al evaluador humano la responsabilidad de evaluar cada uno de los formularios direccionables por la expresión de la figura 4. Lo mismo se aplica para la regla 12.3b de la figura 5 (para opciones no agrupadas dentro de un select).

WCAG #	Regla	HERA	WebXACT,Taw,WAVE
6.3b	No obstruir la navegación con JavaScript	Auto Obj.	Manual
6.4a	El evento `onmouseover` debe emparejar con `onfocus`	Auto Obj.	Manual
6.4b	El evento `onmouseout` debe emparejar con `onblur`	Auto Obj.	Manual

Cuadro 7: Reglas con comportamientos diferentes según la herramienta (HERA frente al resto)

//form[*toomany_inputs*(.//input)][not(.//fieldset)]

Fig 4: Expresión XPath para detectar los formularios que rompen la regla WCAG 12.3a

//select[*toomany_options*(option)][not(optgroup)]

Fig 5: Expresión XPath para detectar los conjuntos de opciones que rompen la regla WCAG 12.3b

4. Conclusiones

A partir de los resultados de la sección 3.3, es fácil no sorprenderse de que los resultados de una evaluación de accesibilidad (para una página dada) sean diferentes según la herramienta utilizada, pues hay un elevado número de páginas Web que pasan los tests automáticos de las herramientas menos exigentes y sin embargo no son capaces de aprobar el test de las herramientas más exigentes. Ciertamente, de comparar las secciones 3.2 y 3.3 se puede desprender la conclusión de que existen más diferencias que puntos en común entre las distintas herramientas.

Es decir, que si bien estas herramientas pueden resultar bastante útiles para alguien inexperto en temas de accesibilidad, pues con casi toda probabilidad detectarán alguna barrera en páginas donde los criterios de accesibilidad no han formado parte del diseño, para un usuario con buenos conocimientos, estas herramientas presentan lagunas importantes que podrían ser fácilmente mejoradas. En este sentido, estamos trabajando en una herramienta de evaluación que, en lugar de comprobar las condiciones de los puntos de chequeo con algoritmos específicamente escritos en un lenguaje de programación convencional, use expresiones XPath [3,4], XPointer [6] y XQuery [5] para determinar las barreras de un sitio Web.

Por otro lado, cabe destacar que tampoco ninguna herramienta es perfecta y que todavía no existe una que pueda reemplazar a todas las demás, pues existen muchas condiciones automatizables que estas herramientas no comprueban. La sección 3.4 ha mostrado ejemplos de comprobaciones aún no cubiertas por ninguna de las herramientas estudiadas y ha proporcionado soluciones para implementarlas fácilmente con reglas preconstruidas en XQuery.

En general, se puede concluir que, si bien es imposible automatizar todos los puntos de chequeo, la mejor labor la puede desarrollar el autor de contenidos Web, especialmente si escoge un formato específicamente orientado a la accesibilidad como XHTML Basic. Este formato mejora la accesibilidad reduciendo el número de comprobaciones manuales y aumentando el número de comprobaciones automatizables.

Agradecimientos

El trabajo en el que se ha basado este artículo ha recibido el apoyo de los proyectos INFOFLEX *TIC2003-07208* y SIEMPRE *TIC2002-03635* financiados por el Programa Nacional Español de Tecnologías de Información y Comunicaciones.

Referencias

[1] W3C *Web Content Accessibility Guidelines 1.0*
`www.w3.org/TR/WCAG10`

[2] W3C *Techniques For Accessibility Evaluation And Repair Tools W3C Working Draft, 26 April 2000*
`www.w3.org/TR/AERT`

[3] W3C *XML Path Language (XPath) Version 1.0 W3C Recommendation 16 November 1999*
`www.w3.org/TR/xpath`

[4] W3C *XML Path Language (XPath) 2.0 W3C Working Draft 29 October 2004*
`www.w3.org/TR/xpath20`

[5] W3C *XQuery 1.0: An XML Query Language W3C Working Draft 29 October 2004*
`www.w3.org/TR/xquery`

[6] W3C *XML Pointer Language (XPointer), W3C Working Draft 16 August 2002*
`www.w3.org/TR/xptr`

[7] W3C *XHTML 1.0 ^{TM} The Extensible HyperText Markup Language (Second Edition), A Reformulation of HTML 4 in XML 1.0, W3C Recommendation 26 January 2000, revised 1 August 2002*
`www.w3.org/TR/xhtml1`

[8] W3C *XHTML Basic W3C Recommendation 19 December 2000*
www.w3.org/TR/xhtml-basic

[9] W3C *XHTML ™ 1.1 - Module-based XHTML, W3C Recommendation 31 May 2001*
www.w3.org/TR/xhtml11

[10] W3C *XHTML ™ 2.0, W3C Working Draft 22 July 2004*
www.w3.org/TR/xhtml2

[11] W3C *Markup Validation Service*
validator.w3.org

[12] W3C *CSS Validation Service*
jigsaw.w3.org/css-validator/

[13] Watchfire *Bobby Accessibility tool*
bobby.watchfire.com/bobby/html/en/index.jsp

[14] CEAPAT, Fundación CTIC, Spanish Ministry of Employment and Social Affairs (IMSERSO) *Online Web accessibility test*
www.tawdis.net

[15] Watchfire *WebXACT*
webxact.watchfire.com

[16] Vicente Luque Centeno, Carlos Delgado Kloos, Luis Sánchez Fernández, Norberto Fernández García *Device independence for Web Wrapper Agents*
Workshop on Device Independent Web Engineering (DIWE'04) 26 July 2004, Munich

[17] Peter Blair *A Review of Free, Online Accessibility Tools*
www.webaim.org/techniques/articles/freetools, February 2004

[18] Sourceforge *JTidy*
jtidy.sourceforge.net

[19] Fondazione Ugo Bordoni *Torquemada, Web for all*
www.webxtutti.it/testa_en.htm

[20] Fundación Sidar *HERA*
www.sidar.org/hera

[21] WebAIM (Web Accessibility in Mind) *WAVE 3.0 Accessibility Tool*
wave.webaim.org/index.jsp

Arquitectura de una Solución de Optimización Lógica y Física de Consultas en Mediadores de Fuentes Web

Justo N. Hidalgo, Alberto Pan, José Losada, Manuel Álvarez

Denodo Technologies
Calle Real, 22 3º
A Coruña 15003, España
Teléfono: +34 981 100 200
E-mail: {jhidalgo, jlosada}@denodo.com

Dpto. de Tecnologías de la Información y la Comunicación
Universidad de A Coruña
Campus de Elviña s/n A Coruña 15071, España
Teléfono: +34 981 167 000
E-mail: {apan, mad}@udc.es

Abstract. *Access to data sources with heterogeneous formats is one of the most important challenges in Enterprise Information Integration. Even though there exists an increasing number of solutions which obtain information from structured and semistructured sources such as web sites, lack of QoS in HTTP connections makes it hard to measure the mediated query performance. Different solutions based on relational and hierarchical optimization techniques have been proposed to improve results in terms of time and space. This paper proposes the utilization of some of these techniques in an integrated way, adding some concepts which help improve their functioning, all by using a specialized cost repository; some of them are the differentiation between possitive and negative query capabilities or the use of search method's utilization ranking in the cost repository. Besides, it explains how physical layer optimization by means of a configurable web navigator pool helps improve query's result time in a non-despicable way.*

1 Introducción

Las bases de datos virtuales o mediadores permiten la obtención, unificación y muestra de datos que se encuentran en entornos heterogéneos, tanto debido a diferentes tipos de formato como a protocolos de acceso y a su almacenamiento en diferentes repositorios geográficamente independientes. Así, una base de datos virtual será capaz de agregar información de una base de datos relacional con información procedente de un entorno semiestructurado, como la web.

A la hora de realizar una consulta en una base de datos virtual, se genera un conjunto de planes de consultas; todos son capaces de responder correctamente a la consulta, pues cumplen sus capacidades y restricciones. La diferencia entre ellos radica en la utilización de diferentes estrategias de operadores, de fuentes complementarias (sitios web que disponen del mismo tipo de información y por tanto pueden responder a una subconsulta concreta de la misma manera), y de métodos de búsqueda de una misma fuente que también pueden resolver la subconsulta. Sin embargo, en un momento concreto, sólo uno de esos planes será el más óptimo, es decir, será el que permita una ejecución más rápida, optimizando los recursos a utilizar. Para poder conseguir ese plan óptimo en cada consulta –o, al menos, aproximarse estadísticamente a ello-, hay que proceder a utilizar modelos de costes que permitan medir cada plan. La importancia que tiene la mejora de tiempos de respuesta en entornos web, donde no se asegura la calidad de servicio, ha impulsado esta línea de investigación en los últimos años.

Tal y como se describe en [7], la estimación de costes en bases de datos relacionales utiliza estadísticas de la base de datos y fórmulas para computar el coste de cada operador en cada plan. Aunque muy útil para entornos locales y/o perfectamente estructurados, este enfoque no es directamente aplicable a las bases de datos virtuales, debido principalmente a las siguientes razones:

- Las fuentes no suelen ofrecer información estadística.
- Las fórmulas de coste para cada operador varían dependiendo de:
    - La implementación del *wrapper* (interfaces a las fuentes remotas).
    - La implementación de la fuente remota.
- Los costes de comunicaciones no se determinan fácilmente y pueden variar.

Por ello, la utilización de costes en entornos heterogéneos no se define como una proyección directa de las investigaciones previas en el mundo relacional. Existen diferentes opciones estudiadas a lo largo de los años. Por ejemplo, en el arriba mencionado artículo, los autores defienden una combinación de un modelo genérico de costes con información específica de cada *wrapper*. De esta manera, el *wrapper* especifica lo que pueda, y, de lo que no sepa, ya se encarga el mediador con la información por defecto. Tal y como está concebido, este concepto es muy adecuado cuando la información es muy homogénea entre sí.

Desgraciadamente, no se tiene en cuenta el coste menos medible de todos, que es el de comunicaciones, lo cual lo hace inviable en un entorno real, pues la capacidad genérica del mediador puede no sólo ser inadecuada, sino que puede empeorar los tiempos de respuesta [3] [8].

En [13] se es más extenso en cuanto al grupo de parámetros necesarios contando con el coste de comunicaciones entre los agentes y las fuentes remotas. Sin embargo, las fórmulas de propagación de costes que provee no son lo suficientemente completas como para ajustarse a un sistema real.

Otras investigaciones han mejorado u ofrecido alternativas a estos trabajos en diferentes ámbitos. [4] se basa en la estimación de los coeficientes de un modelo de costes genérico. Esta aproximación no suele funcionar en entornos heterogéneos, pues el modelo cambia en cada fuente. De hecho, al no permitir sobrecarga por parte de los *wrappers*, es incluso menos apropiado para mediadores de este tipo.

Una investigación que tiene más en cuenta los factores intrínsecos de la mediación heterogénea es el proyecto Hermes, cuya solución se encuentra especificada en [1]. En ella, el modelo de costes se evalúa a partir de información histórica (almacena la información de costes de cada consulta anterior sobre cada fuente remota), lo que produce una caché de estadísticas de invocaciones reales, para estimación posterior del coste de planes de ejecución. Esta postura tiene dos consecuencias directas:

- Por una parte, esta medición histórica resulta muy útil para fuentes utilizadas uniformemente.

- Sin embargo, resulta inútil para fuentes poco utilizadas o para fuentes con predicados poco utilizados. Obviamente, ya que esta técnica se basa en históricos, es necesario que existan datos precedentes.

Aunque en muchos casos las fuentes vayan a ser utilizadas uniformemente (los buscadores de productos sobre todas las fuentes remotas disponibles, aunque tengan diversas capacidades de consulta, serán accedidas uniformemente) es importante tener en cuenta la segunda consecuencia. La opción de inicializar la información con datos por defecto altos puede llevar a que un método de búsqueda no se ejecute nunca, ya que, aunque su coste real sea el más bajo, se ha elegido otro desde el principio cuyo coste es menor que el "default" del primero. La de inicializar con datos bajos (tiempo = 0) soluciona ese problema, afectando sólo las primeras consultas (cada consulta subsiguiente irá accediendo a un método de búsqueda no inicializado, pues su valor será 0, siempre menor que los del resto).

El concepto básico de Hermes ha sido utilizado y mejorado por otros autores, como [14], donde se detalla el concepto de optimizador en dos fases, o [5], con sus estudios sobre adaptación dinámica de consultas a partir de los costes obtenidos. Sin embargo, ninguna de las investigaciones hasta el momento ha pretendido agregar de manera consistente las diferentes herramientas y algoritmos de optimización en un sistema único. Este articulo presenta una arquitectura unificada de optimización de consultas basado en información histórica almacenada en un repositorio de costes, mediante una fórmula de procesamiento estadístico que optimiza la decisión de qué plan de consulta seleccionar.

Por otra parte y como punto fundamental, este artículo propone la utilización de técnicas de optimización de la capa física de los mediadores, pretendiendo aumentar en órdenes de magnitud las prestaciones de estos sistemas con respecto a las alternativas de optimización tradicionales.

2 Análisis del Problema

En nuestra experiencia [8], sin duda el factor que influye más decisivamente en los tiempos de respuesta es la transmisión de información a través de la red. En particular, las fuentes web presentan los siguientes problemas fundamentales:

- En fuentes con mantenimiento de sesión normalmente es necesario realizar varias peticiones de páginas (secuencias de navegación) para hacer la consulta. También puede ser necesario (y normalmente lo es) navegar a través de varias páginas para obtener todos los resultados a una consulta (respuestas en intervalos).

- El formato HTML tiene muchas veces un altísimo *overhead*. No es nada raro que una página que contenga 10 tuplas de la respuesta a una consulta (4-5 Kbytes útiles) tenga un HTML que ocupe 40-50 Kbytes.

- No se puede asegurar un nivel de Calidad de Servicio en ninguna transmisión a través de la World Wide Web, debido a que el protocolo de comunicaciones subyacente, TCP/IP, no lo permite.

Si bien es verdad que en fuentes de otros tipos el problema de la transmisión de datos es menor, ya que normalmente se requiere una sola conexión por consulta y el ratio de *overhead* será mínimo, debe tenerse en cuenta que si bien no siempre todas las fuentes serán web, en casi todas las aplicaciones habrá al menos alguna fuente de este tipo (ya que, por un lado, es un tipo de fuente muy importante y, por otro, la utilización de Bases de Datos Virtuales aumenta con respecto a otros enfoques, como el Data

Warehouse, en estos casos). Si esa fuente o fuentes web están involucradas en alguna consulta (especialmente en las que sean inherentemente síncronas) retrasarán la ejecución de toda la consulta, con lo cuál el cuello de botella seguirá siendo el mismo.

Por otro lado, aún cuando el sistema no acceda a fuentes web, el parámetro más costoso seguirá siendo la transmisión por la red, aunque la diferencia sea menos acusada.

Nuestra propuesta se basa en mejorar la optimización de la capa lógica, mediante la utilización de diferentes mecanismos integrados de control de estado de la información estadística, así como de caché; por otra parte, el artículo detalla la utilización de un sistema de optimización de la capa física del sistema mediador, aportación novedosa y que aporta mejoras de rendimiento órdenes de magnitud mayores en casos óptimos.

3 Arquitectura del Sistema Mediador

En la Fig. 1 se muestra la arquitectura de un sistema mediador. El mediador recibe consultas sobre el esquema mediado expresadas en algún lenguaje de consulta, como SQL u OQL.

Una vez recibida la consulta, el Generador de Planes genera los planes de ejecución alternativos de la misma. Cada plan estará compuesto por un conjunto de subconsultas a realizar sobre un conjunto de fuentes; cada uno de estas fuentes devolverá una serie de resultados que, tras ser postprocesados y unidos, permiten obtener el resultado final.

El Generador de Planes puede generar gran cantidad de posibles planes de consulta para una única consulta por parte del usuario, debido a que podrá tener en cuenta alternativas tales como diferentes estrategias de ejecución de operadores, fuentes de datos complementarias, o métodos de búsqueda de una misma fuente que respondan a la misma subconsulta. Es tarea del Optimizador Lógico el elegir el plan de consulta óptimo para este momento concreto, que es el que será enviado al Motor de Ejecución.

Por otra parte, el w*rapper* es el elemento de la arquitectura que emitirá la subconsulta deseada utilizando el formato y el protocolo que la fuente remota acepte. En el caso de fuentes web, los *wrappers* son navegadores web capaces de comportarse de la misma manera que un usuario humano, que navega a través de la jerarquía del dominio web. Estos navegadores son creados y gestionados a través de un *pool* de navegadores. Este *pool* permite, además, monitorizar y controlar la optimización física que se comenta en un apartado posterior.

Resumiendo, a partir de esta arquitectura, la optimización de las consultas debe verse desde dos puntos de vista diferentes: por una parte, la elección del plan de consulta óptimo a partir de la consulta realizada por el usuario, pero, por otra parte, la optimización del acceso a esas fuentes web. La optimización del plan de consulta se realiza en la capa lógica del mediador, mientras que la optimización del acceso a las fuentes web se hace en la física. Muchos trabajos anteriores [6][7][9][12][14] se han centrado en la optimización de consultas, basándose en las dificultades encontradas en las bases de datos jerárquicas y relacionales. Sin embargo, y debido a que el punto crítico en el acceso a fuentes remotas es el tiempo de comunicación, es necesario prestar mayor atención a cómo acceder de una manera más eficiente a los recursos.

4 Optimización de Capa Lógica

Los tipos de optimización que se plantean en este artículo se detallan a continuación.

4.1 Delegación de Trabajo a las Fuentes

Aunque es posible encontrar ejemplos teóricos en los cuáles, en ciertas condiciones, un plan que delega todo el procesamiento posible a las fuentes remotas sea menos eficiente que un plan que realice post-procesamiento, en general nunca será así, y menos con consultas que involucren fuentes web. Esta heurística se asume en prácticamente todos los mediadores [2].

En principio, uno podría querer delegar a las fuentes cualquier tipo de operación que permitan sus capacidades de consulta: selecciones, proyecciones, *joins* y uniones. Aunque cualquier delegación permitirá una optimización mayor del tiempo de respuesta, la delegación de selecciones es la más común y la que, estadísticamente, mejor comportamiento conlleva [4]. Delegar uniones y *joins* de forma dinámica en un mediador es mucho más complejo de construir (definir estáticamente relaciones base cuyas tuplas se obtengan a partir de *joins*, uniones o cualquier otra operación en la fuente es algo ya conseguido en la plataforma de estudio [10]), y su importancia es relativa:

Delegar <A *join* B> permite disminuir la cantidad de datos a transmitir ya que los atributos de *join* se transmiten una sola vez, mientras que si transmitimos A y B por separado y hacemos el *join* en el mediador, vendrán dos veces. Sin embargo, propagar las selecciones decrementa el volumen de datos bastante más.

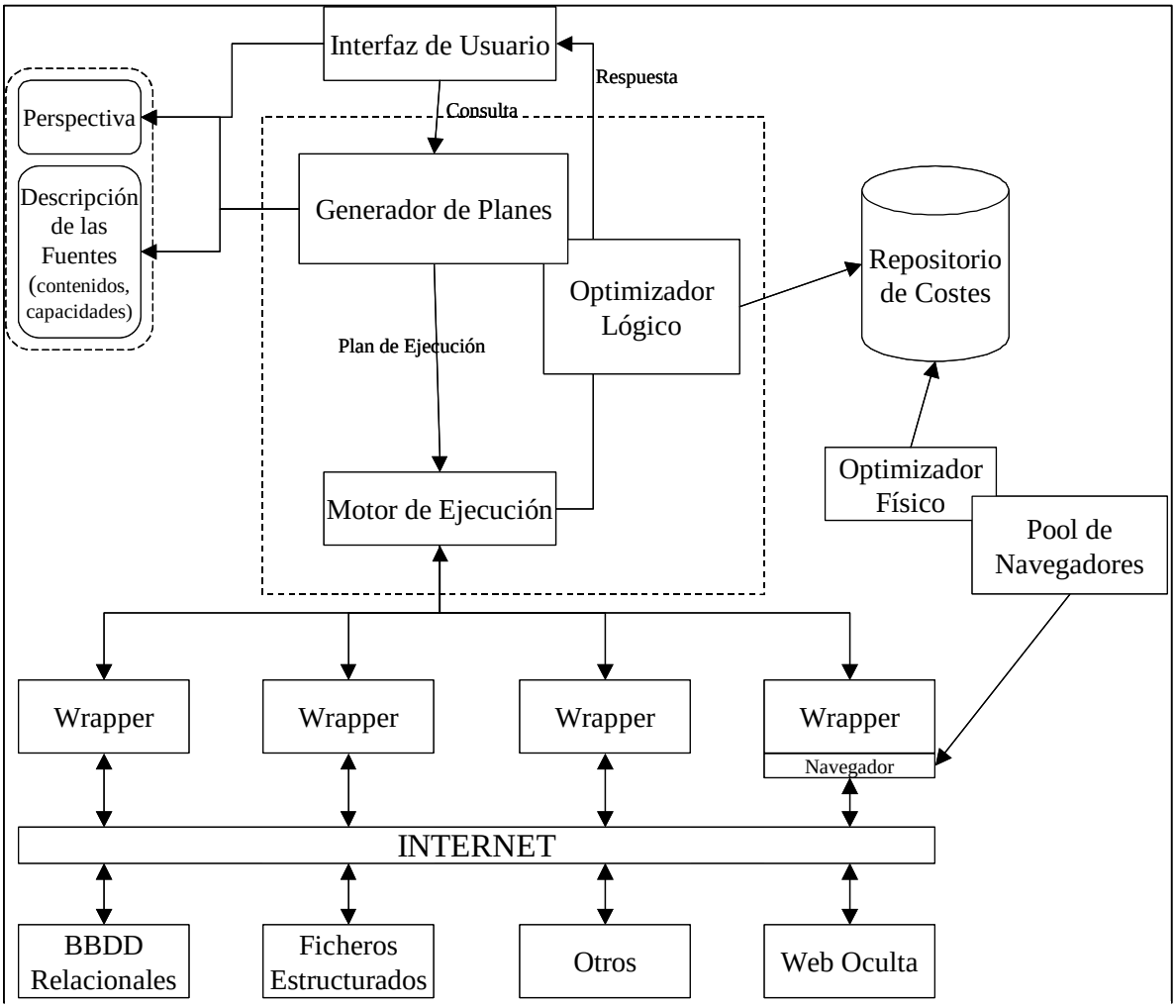


Figura 1: Arquitectura de una Base de Datos Virtual con Optimización Lógica y Física

Puede parecer en principio que delegar uniones NO permite disminuir la cantidad de datos a transmitir, ya que efectivamente, delegar <A Union B> no disminuye la cantidad de datos a transmitir con respecto a traerse A y B por separado y hacer la unión en local. Sin embargo, si se pudiesen delegar *joins*, pero no uniones, entonces no podría delegarse una consulta como <A Union B> *join* C que sí podría delegarse si el mediador delegase también uniones, con lo que, después de todo, sí sería bueno también poder delegar uniones.

Sólo las fuentes de tipo base de datos relacional permitirán normalmente *joins* y uniones. E incluso en estos casos, la mayoría de las veces la cosa podrá solucionarse aceptablemente definiendo algunas consultas patrón estáticas mediante *wrappers* JDBC/ODBC.

Por tanto, aunque sería bueno delegar *joins* y uniones, las ventajas obtenidas no parecen demasiado grandes con respecto al coste de diseño e implementación. Como se explicará en el apartado 5, otros tipos de optimizaciones pueden obtener una mejora más radical en el entorno distribuido en el que se encuentra un sistema mediador.

4.2. Elección de Estrategia de Ejecución de Operadores en la Operación *Join*

Cada una de las operaciones de álgebra relacional puede tener diferentes estrategias de ejecución. En concreto, nuestra plataforma escoge entre *Join* Normal y *Join* Anidado.

Por experiencia, normalmente compensará casi siempre el *join* normal cuando sea posible, ya que las latencias asociadas a emitir una nueva consulta suelen ser muy altas en comparación con la transmisión de resultados por la red (aunque puedan ser más resultados a transmitir en un *join* normal que en un *join* anidado). Sin embargo, uno puede pensar fácilmente casos en los que no es así. Por ejemplo, si las consultas involucradas en el *join* normal dan tantos resultados que obligan a paginar mucho, mientras que las del *join* anidado dan muy pocos, el número de conexiones debidas a la paginación podría superar al número de conexiones debido a la estrategia de *join* anidado.

Para saber si compensa hacer un *join* normal o un *join* anidado, se plantea la utilización de estadísticas de cardinalidad y costes de los métodos de búsqueda

para cada relación base. La obtención se realiza mediante un sistema basado en [1], al cuál se añaden nuevas capacidades tales como la utilización de intervalos horarios de utilización del sistema, que permite una mayor flexibilidad de optimización y un mayor control de granularidad de la información de costes, lo que redunda en una mayor calidad estadística.

4.3. Elección de Método de Acceso

La gran cantidad de fuentes disponibles en la World Wide Web plantea la posibilidad de disponer de fuentes redundantes, es decir, de webs con información solapada, de manera que una u otra pueden ser invocadas para obtener resultados idénticos. Sitios de información bursátil, de cambio de moneda o de información meteorológica son buenos ejemplos de este tipo de webs.

El sistema, tras asociar un coste a cada uno de estos métodos de búsqueda alternativos, escogería en cada momento el de menor coste cuando la consulta pudiese ser efectuada mediante ambos métodos; por otra parte, esto nos permite implementar fácilmente mecanismos de control de errores mediante reintentos: si uno de los métodos de búsqueda no responde, antes de devolver un error declarando la imposibilidad de ejecutar el plan de consulta al que pertenece, se intentará ejecutar cada uno de los métodos de búsqueda alternativos. Resumiendo, este esquema nos permite elegir dinámicamente de qué fuente extraer los datos en cada momento en función de la consulta efectuada, del tramo horario, etc.

El apartado 6 detallará la información almacenada en el repositorio de costes, así como su operativa de funcionamiento.

5 Optimización de Capa Física

Hay fuentes web complejas que requieren secuencias de navegación largas antes de llegar al formulario real de consulta. Esto es debido a los sistemas de mantenimiento de sesión. Por ejemplo, en muchas fuentes reales hay que introducir un par usuario/contraseña y después se tiene que navegar a través de una o dos páginas hasta llegar a un determinado formulario de consulta. Esto hace que una consulta pueda llevar tiempos muy largos.

El flujo de navegación de muchas consultas tienen "denominadores comunes", elementos que se repiten, tales como el acceso mediante usuario/contraseña, o las diferentes opciones por las que se navega antes de llegar a elementos concretos de una fuente web. Se pierde mucho tiempo de procesamiento debido a que el sistema no tiene en cuenta que de una consulta a otra sólo sería necesario un paso atrás y una nueva selección, sino que repite el flujo completo de navegación desde el principio.

La optimización del flujo de navegación puede proporcionar mejoras muy considerables en cuanto al tiempo de acceso a la información remota de cada fuente, llegando a mejorarlo en órdenes de magnitud –sobre todo en fuentes residentes en servidores con poca capacidad de ancho de banda, o en fuentes con muchos pasos intermedios-. El sistema mediador presentado permite sistematizar el acceso a fuentes web transaccionales –aquellas que requieren navegación a través de formularios de búsqueda o autenticación, menús, enlaces, etc.- mediante la utilización de un lenguaje de secuencias de navegación (NSEQL: Navigation Sequence Query Language) que es utilizado por navegadores automáticos que emulan el comportamiento de un usuario. Estos navegadores son recursos críticos controlados por un *pool* de navegadores, que gestiona el ciclo de vida de estos componentes. La solución propuesta en este artículo es la extensión del *pool* de navegadores, de manera que se almacena, para cada navegador del *pool*, la secuencia que le llevó a su estado actual. Cuando se recibe una nueva secuencia para su ejecución, el *pool* comprueba si hay algún navegador que tenga alguna secuencia con un prefijo común (es decir, con un subconjunto de esa secuencia desde el principio de ésta) con la nueva secuencia. En ese caso, se evalúa si es mejor reutilizar el estado actual del navegador o bien empezar desde el principio. Así, en el caso comentado anteriormente, al llegar la secuencia de la segunda consulta, el *pool* se daría cuenta de que ya hay un navegador que está a un solo paso de completarla y lo reaprovecharía.

De hecho, y para evitar el problema del coste de inicialización (i.e. el coste derivado de que un navegador determinado tenga que alcanzar el estado de espera por primera vez), el *pool* de navegadores, al arrancar, pone en marcha un número preconfigurado de navegadores ya inicializados, es decir, que se encuentran en los estados –en las páginas concretas- más adecuados para responder las consultas que se prevé que vayan a llegar. Esto se consigue mediante un histórico de ranking de fuentes y estados que también se almacena en el repositorio de costes, de manera que podemos construir un modelo heurístico.

El *pool* de navegadores ha de disponer tanto de información de estado como estadístico de costes o, más exactamente, de accesos.

- La información de estado es la que utilizará el *pool* para saber dónde están sus navegadores en cada momento. Actualmente, la información que se maneja es el conjunto de secuencias de navegación, donde se define secuencialmente cuál es el camino que debe seguir el navegador. Cada elemento de esta secuencia es realmente un estado por el que pasa el autómata finito determinista que es. Esta información es la que el *pool* necesita conocer de cada elemento de

manera que, cuando llega una nueva consulta, pueda determinar cuál de sus elementos tiene una distancia menor con respecto al objetivo –última navegación- de la nueva secuencia a ejecutar.

- La información estadística de accesos es la que permitirá al *pool* inicializar sus navegadores de manera que, estadísticamente, las siguientes secuencias sean lo más próximas posibles a las primeras. Cuanto más próximas, el valor de coste será menor.

5.1 Algoritmo de Optimización Física del *Pool* de Navegadores

La propagación de costes en tiempo de preejecución tiene como inconveniente la imposibilidad de contar con información dinámica que influye en gran medida en la optimización real. Como se ha comentado anteriormente, la propuesta de este artículo es la utilización de información histórica adecuadamente procesada que ayude al generador de consultas a seleccionar más adecuadamente.

Cada vez que se realice una consulta con una secuencia determinada, el repositorio de costes debe almacenar ese hecho, de manera que se vaya creando un ranking de secuencias por tramo horario. Esto permitirá que el algoritmo de esta capa sea el siguiente: en primer lugar, el *pool* de navegación, nada más arrancar, arrancará el número de navegadores indicados en el fichero de configuración. Cada uno de estos navegadores utilizará directamente el conjunto de secuencias mejor posicionadas en el ranking -no tiene porqué ser un navegador por cada secuencia, sino que si el número 1 del ranking tiene muchísimas más ocurrencias que el resto, dos o más navegadores utilizarán ésta; para ello se utiliza una distribución de Poisson-. Estos navegadores se pararán en el estado de la secuencia en la que sea necesario introducir información en tiempo de ejecución –usuario, información de entrada, etc.-; de esta manera, se está optimizando al máximo la utilización de éstos.

Cuando se lanza un nuevo navegador, hay que contar con dos tipos de información: la provista por el repositorio de costes, más la información que provea el *pool* acerca de los navegadores que ya se encuentran en activo –de manera que, por ejemplo, si los navegadores en activo ya copan las dos primeras secuencias del ranking, el nuevo navegador vaya a la tercera secuencia-.

Al llegar una petición que implique utilizar un navegador, el *pool* comprobará si existe alguno con la misma secuencia. Si recibe más de un resultado, decide cuál tiene que realizar menos secuencias para llegar al resultado requerido –estas acciones se implementan en un sólo paso, utilizando el concepto de distancia-.

Por otra parte, puede ocurrir –y ocurrirá- que la secuencia no tenga par en los navegadores activos (i.e. ningún navegador se encuentra en ninguna página perteneciente a la secuencia deseada). Esta situación puede llevar a varias posibilidades:

- Si la probabilidad de que esta secuencia vuelva a repetirse es MUCHO menor –valor parametrizable- que las secuencias que ya se encuentran en activo, puede ser planteable el lanzar un nuevo navegador aunque el resto no se esté utilizando, ya que si se plantea la otra opción, es posible que lleguen nuevas consultas que sí las utilicen.

- Siguiendo con el caso anterior, puede que el *pool* no acepte la creación de un nuevo navegador –porque ya haya un máximo en activo-. En este caso, la secuencia utilizará el navegador con la secuencia de menor probabilidad de ocurrencia.

Esta será la misma manera de actuar en caso de que las diferencia entre probabilidades no sea muy alta, o incluso sea negativa.

Al igual que el almacenamiento de información de costes por parte del *wrapper*, este repositorio habrá de utilizar una fórmula estadística que tenga en cuenta el estado y el tramo horario, de manera que no sea simplemente una media, sino que tenga en cuenta la variación temporal entre valores.

5.2 Algoritmo de Vuelta Atrás a Página Común

Una vez que el navegador ha devuelto la información requerida por la secuencia de navegación, navegará de vuelta al elemento de secuencia más favorable para la siguiente petición. Si la siguiente petición proviene de otro usuario, la página anterior a la autentificación es el mejor sitio a donde ir; sin embargo, si proviene de un mismo usuario, puede que le merezca la pena quedarse en otra URL, donde la autentificación ya se haya producido.

Aunque todavía es necesario un estudio más detallado, el enfoque utilizado en la actualidad es el siguiente. El navegador, tras finalizar la ejecución de una secuencia, pregunta al gestor de costes del *pool* cuál es la consulta concreta de esta secuencia –es decir, secuencia con datos concretos- con el ranking más alto. El navegador se dirigirá hasta la intersección de ambas secuencias; esta página intersección será:

- O la misma secuencia concreta anterior (o quizá con diferentes parámetros de consulta en el último elemento que lo permita).

- O la última página de la secuencia con *ranking* más alto.

De esta manera, se intenta coordinar el aspecto estadístico de la solución con el heurístico. Aunque estadísticamente es el número 1 del ranking el que aparecerá con más probabilidad, no es menos cierta la heurística planteada.

5.3 Mantenimiento de sesión en los navegadores

Un problema que surge aquí es que la sesión de los navegadores puede caducar haciendo que al intentar reutilizarlos, la secuencia falle. Obviamente, mantener durante un tiempo indefinido un navegador en una URL particular no es posible. Para solucionarlo hay varias opciones: una es hacer que los navegadores accedan aleatoriamente a un link y después hagan una vuelta atrás a ciertos intervalos. Otro es sencillamente que si la secuencia falla, entonces se rehaga desde el principio. La opción elegida es la parametrización del *session timeout* de cada fuente web –con un valor por defecto-, para que el *pool* de navegadores actúe antes de que la sesión caduque.

6 Repositorio de Costes

Como se ha podido comprobar a lo largo de este artículo, la optimización de consultas en un mediador heterogéneo requiere de la utilización constante de información histórica. El repositorio de costes es el elemento que almacena esa información, estructurándola de manera que los módulos de optimización física y lógica los usen fácilmente.

La tabla 1 muestra los componentes básicos del repositorio de costes (no aparecen los valores agregados procesados a partir de fórmulas estadísticas). Para cada fuente y método de búsqueda, se almacena tanto la lista de atributos de las capacidades negativas como la de las positivas. Las capacidades negativas se definen como las restricciones que deben cumplir las consultas recibidas por una fuente para que ésta sea capaz de ejecutarlas, mientras que las positivas representan las consultas que puede ejecutar la fuente por sí misma una vez que las restricciones descritas por las capacidades negativas han sido satisfechas [10]. La división en el repositorio de costes se realiza para poder tener información acerca del coste asociado a los postprocesados (capacidades de consulta que no aparecen como positivas).

El repositorio almacena los valores de parámetros espaciales y temporales de cada método de búsqueda. La división entre tiempo de conexión, transición y el resto, permite que el repositorio pueda guardar información global para varios métodos de búsqueda –porque, por ejemplo, su flujo de acceso se realice a través de la misma URL, por lo que el modo de autentificación y navegación sea el mismo-.

Otro punto importante en el repositorio mostrado es el almacenamiento de información por cada método de búsqueda y tramo horario. A través de estudios realizados [3], se observan las grandes diferencias de tiempo existentes en diferentes tramos horarios. La elección entre un método de búsqueda u otro también depende del tramo horario en el que se realiza la consulta. Por último, los datos sobre porcentajes de conexiones fallidas permitirán al mediador poder elegir entre fuentes rápidas, pero proclives a fallar, y fuentes algo más lentas, pero más fiables.

Por otra parte, el repositorio de costes también almacena información necesaria para la optimización física, como la ruta de acceso completa para cada método de búsqueda, y el ranking de utilización, el cuál permitirá al *pool* de navegadores seleccionar con qué ruta de acceso inicializar los nuevos navegadores.

Tabla 1: Repositorio de Costes

TIPO DE OPTIMIZACIÓN	PARÁMETRO
LÓGICA	Nombre de Fuente
	Método de Búsqueda
	Lista de Atributos (capacidades positivas y negativas): - Clave - Valor con multiplicidad - Operador de la consulta
	Parámetros Temporales: - Tiempo de inicialización - Tiempo de devolución de primera lista de resultados - Tiempo de devolución de siguientes resultados
	Parámetros Espaciales: - Número de tuplas en primera página - Número total de tuplas - Tamaño de primera página
	Tramo temporal
	Porcentaje de conexiones fallidas
LÓGICA (CACHÈ)	Lista de consultas con y sin cachè
	Número de consultas en cachè
	Número total de consultas
	Tiempo de acceso a cachè –en caso de cachè distribuida-
FÍSICA (*POOL*)	Ruta completa de acceso
	Lista de accesos al *pool*: - Fecha - Subconjunto de rutas de acceso

Para cada método de búsqueda, el repositorio de costes guarda estos datos de dos maneras diferentes, basándonos en la idea original de [1] de reducciones mediante tablas *lossy* frente a tablas *lossless*. Por una parte, se almacenan los datos estadísticos de cada

consulta a una fuente realizada por un *wrapper*. Esto nos permite disponer en todo momento de los elementos atómicos que se agrupan mediante diferentes fórmulas estadísticas (medias, varianzas, o fórmulas de estado) para también disponer de información útil para la toma de decisiones ante un conjunto de planes de consulta. Este repositorio se almacena en una base de datos relacional, de manera que su acceso sea rápido y eficiente.

7 Conclusiones

Este artículo ha descrito una arquitectura de optimización de consultas sobre bases de datos virtuales. El sistema aprovecha un conjunto de ideas propuestas por diferentes autores en los últimos años, pero ofreciendo al mismo tiempo elementos innovadores inexistentes hasta ahora en la literatura, como diferenciar entre capacidades positivas y negativas para mejorar las prestaciones en los post-procesados y, por encima de todo, la optimización de la capa física que, en la mayor parte de las ocasiones, permite que el sistema mejore sus prestaciones en órdenes de magnitud. Las características expuestas en este artículo se encuentran implementadas en la herramienta software Virtual DataPort de Denodo Technologies [11].

El trabajo futuro a realizar sobre la optimización de la capa lógica se centra en investigar algoritmos que permitan la automatización completa del proceso de toma de decisión mediante la información provista por el administrador o usuario en tiempo de ejecución. La mejora de la ordenación de la operación de *join* basándose en enfoques relacionales es otro tipo de optimización que será integrada en el sistema actual.

En cuanto a la optimización de la capa física, queda todavía mucho camino por recorrer. Un paso a realizar a corto plazo es que el sistema "aprenda" de los *session timeouts.* Cada vez que se produzca un timeout, la fuente navega automáticamente al elemento que lo provocó, y modifica el valor del parámetro de *timeout*, reduciéndolo (p.e. x' = x/2).

Referencias

[1] Adali, S., Candan, K., Papakonstantinou, Y., Subrahmanian, V.S. Query Caching and Optimization in distributed mediator systems. In Proceedings of the ACM SIGMOD Conference on Management of Data. 1996

[2] BEA Liquid Data for WebLogic 8.1 DataSheet: http://www.bea.com/content/news_events/white_papers/BEA_LD_for_WL_ds.pdf

[3] Carneiro, V., Hidalgo, J., Mato J., Orjales, V. Gestión de Calidad de Servicio en Aplicaciones Web. In Tercer Congreso Iberoamericano de Telemática CITA'2003. 2003

[4] Du, W., Krishnamurthy, R., Shan, M. Query Optimization in a Heterogeneous DMBS. En Proceedings of the 18th International Conference on Very Large Databases. 1992

[5] Garcia-Molina, H., Ullman, J.D., Widom, J. Database Systems. The Complete Book. Ed. Prentice Hall. ISBN: 0-13-031995-3

[6] Ives, Z.G., Halevy, A. H., Weld D. S. Adapting to Source Properties in Processing Data Integration Queries. In Proceedings of the ACM SIGMOD Conference on Management of Data. 2004

[7] Naacke, H., Gardarin, G., Tomasic, A. Leveraging mediator cost models with heterogeneous data sources. In Proceedings of ICDE. 1998

[8] Orjales, V., Hidalgo, J., López, G., Carneiro, V. QoS-Meter: Monitorización No Intrusiva de Sitios Web en el Web Oculto. En las IV Jornadas de Ingeniería Telemática, JITEL. 2003.

[9] Ozcan, R., Haas, L. Cost Models DO Matter: Providing Cost Information for Diverse Data Sources in a Federated System. In Proceedings of The VLDB Conference. 1999

[10] Pan, A. Un Sistema Mediador para la Integración de Datos Estructurados y Semiestructurados. Doctoral Thesis. University of A Coruña, Spain. 2002

[11] Pan, A., Raposo, J., Alvarez, M., Montoto, P., Orjales, V., Hidalgo, J., Ardao, L., Molano, A., Viña, A. The DENODO Data Integration Platform. In 28th International Conference on Very Large Databases (SIGMOD VLDB). 2002

[12] Papakonstantinou, Y., Gupta, A., Haas, L. Capabilities-Based Query Rewriting in Mediator Systems. En Proceedings of the Fourth International Conference on Parallel and Distributed Information Systems. 1996

[13] Roth, M.T., Ozcan, F:, Haas, L. M. Cost Models DO Matter: Providing Cost Information for Diverse Data Sources in a Federated System. En Proceedings of the 25th International Conference on Very Large Data Bases, 1999

[14] Zadorozhny, V., Rashchid, L., Vidal, M.E. Efficient Evaluation of Queries in a Mediator for WebSources. In Proceedings of the ACM SIGMOD Conference on Management of data. 2004

Análisis de Arquitecturas basadas en Clusters para Motores de Búsqueda en el Web

Fidel Cacheda, Francisco Puentes, Víctor Carneiro
Departamento de Tecnologías de la Información y las Comunicaciones.
Universidad de A Coruña. Fac. de Informática, Campus de Elviña s/n
15.071 – A Coruña
E-mail: {fidel, fpuentes, viccar}@udc.es

***Abstract**. The increasing number of documents to be indexed in many environments (Web, intranets, digital libraries) and the limitations of a single centralised index (lack of scalability, server overloading and failures), lead to the use of distributed information retrieval systems to efficiently search and locate the desired information. In this work we provide a detailed analysis of the distributed architecture for distributed information retrieval systems, with an special attention to the interconnection network. We analyse the effectiveness of a clustered architecture simulating a variable number of workstations (from 1 up to 4096). A collection of approximately 94 million documents and 1 terabyte (TB) of text is simulated to test the performance of the different architectures. We demonstrate that a clustered system will not outperform a replicated system, due to the reduction of the network bottleneck and moreover a change in the distribution of the users' queries could reduce the performance of a clustered system.*

1 Introducción

Hoy en día se presentan una gran cantidad de retos a los sistemas de Recuperación de Información (IR) en la Web, teniendo en cuenta el dinamismo de este entorno. Uno de los mayores retos es cómo encontrar la información deseada entre todos los datos disponibles. Recientes investigaciones en Web IR han permitido la aparición de motores de búsqueda altamente efectivos que permiten a los usuarios localizar documentos relevantes o útiles. Un segundo reto se centra en cómo diseñar los motores de búsqueda para poder procesar un número masivo de documentos.

Los sistemas de IR basados en un único índice centralizado presentan problemas de falta de escalabilidad [7], reduciendo así su utilidad en colecciones de documentos extremadamente grandes y ante cargas de usuarios elevadas. Por este motivo, se considera más apropiado el uso de métodos basados en Recuperación de Información Distribuida (DIR) para la búsqueda y almacenamiento de información.

Se han definido dos estrategias básicas para la distribución de un índice invertido sobre una colección de servidores de consulta: ficheros invertidos globales y ficheros invertidos locales [14] [17]. En la estrategia de ficheros invertidos globales, cada servidor de consulta almacena una lista invertida correspondiente a una parte de los términos indexados en la colección. En este caso, únicamente algunos servidores de consulta recibirán algunos términos de búsqueda y cada uno devolverá una lista de documentos relevantes para cada término. En la estrategia de ficheros invertidos locales, cada servidor de consulta es responsable de un conjunto disjunto de documentos y, por lo tanto dispone de un índice local independiente. En este modelo, un término de búsqueda es difundido a todos los servidores de consulta, y cada uno responderá con una lista disjunta de documentos relevantes. Tomasic y García-Molina en [17] probaron que la estrategia de ficheros invertidos locales utiliza los recursos del sistema eficientemente, proporciona un buen rendimiento de las consultas en la mayoría de los casos y es más resistente a fallos.

Este estudio es una continuación de nuestro trabajo previo sobre diferentes arquitecturas para sistemas distribuidos de IR introducido en [2] y extendido en [3]. La colección SPIRIT (94.552.870 documentos y 1 terabyte (TB) de texto) [8] fue utilizada en nuestro estudio previo para simular un sistema distribuido básico utilizando la estrategia de ficheros invertidos locales, con el objetivo de analizar el rendimiento para diferentes configuraciones (sistemas distribuidos, replicados y un sistema basado en clusters simple).

En este estudio continuamos haciendo uso de la colección SPIRIT y de la estrategia de ficheros invertidos locales con el objetivo de extender y mejorar diferentes aspectos de nuestro trabajo previo. La primera mejora se centra en la red de interconexión del sistema distribuido, definiendo una red conmutada para analizar las mejoras en el rendimiento frente a una red de acceso compartido. También proporcionamos un análisis más detallado del rendimiento de un sistema de IR basado en clusters frente a un sistema replicado, considerando los cambios de tendencias en las consultas a lo largo del tiempo, basándonos en el trabajo en este área descrito en [15].

Empezamos presentando los trabajos relacionados en la Sección 2. En la Sección 3 describimos nuestro modelo de simulación, centrándonos principalmente en los cambios introducidos en el modelo de la red. En la Sección 4 describimos las simulaciones realizadas para el sistema basado en clusters. Finalmente, presentamos las conclusiones y el trabajo futuro en la Sección 5.

2 Trabajos relacionados

Los trabajos relacionados con los sistemas distribuidos de IR incluyen la evaluación del rendimiento de la arquitectura, la división de los datos, caching y los sistemas multiprocesador. Sin embargo, los trabajos sobre el rendimiento de las arquitecturas son los más directamente relacionados con este artículo.

Harman et al. [5] demuestran la viabilidad de un sistema de IR distribuido, pero sin estudiar aspectos de eficiencia y utilizando una colección pequeña (menos de 1 GB de datos). Desarrollan una arquitectura prototipo y realizan pruebas de usuario, aunque no analizan la eficiencia del sistema.

Burkowski en [1] describe un estudio de simulación que mide el rendimiento de un sistema distribuido de IR utilizando una colección de documentos reducida. Los experimentos realizados exploran dos estrategias para la distribución de la carga a través de los servidores. La primera distribuye la carga uniformemente sobre el conjunto de servidores. En la segunda estrategia los servidores se dividen en dos grupos: evaluación de consultas y recuperación de documentos. Burkowski concluye que la segunda estrategia puede proporcionar mejores tiempos de respuesta bajo ciertas condiciones, aunque la estrategia uniforme se comportará generalmente mejor.

Lin y Zhou [9] implementan un sistema distribuido de IR sobre una red de estaciones de trabajo, demostrando grandes mejoras gracias al paralelismo. Su modelo de recuperación utiliza una variación del esquema de ficheros de firmas para codificar documentos y para distribuir el fichero de firmas a través de la red.

Couvreur et al. [6] analizan el rendimiento de los sistemas paralelos para la búsqueda en grandes colecciones de texto (de más de 100 GB). Se basan en modelos de simulación para investigar tres diferentes arquitecturas hardware (un mainframe, un conjunto de procesadores RISC y una máquina de propósito específico). Su trabajo se centra en analizar el balance entre rendimiento y coste, en donde el mainframe es el más efectivo.

Hawking [6] diseña e implementa un sistema de IR paralelo sobre una colección de estaciones de trabajo, realizando experimentos con un máximo de 64 estaciones. La arquitectura básica del sistema implementado se basa en un proceso central para analizar los comandos de usuario y distribuirlos a los servidores. El proceso central también deberá combinar los resultados parciales antes de mostrar los resultados finales al usuario.

Cahoon y McKinley en [4] describen los resultados de varios experimentos de simulación sobre la arquitectura distribuida INQUERY. Utilizando el comportamiento observado para una implementación mono-servidora, los autores derivan el rendimiento para un sistema distribuido demostrando su escalabilidad. Realizan experimentos con colecciones de hasta 128 GB, demostrando que una arquitectura distribuida simple puede obtener buenos rendimientos para grandes configuraciones.

En [11], Lu y McKinley analizan los efectos de una replicación parcial de la colección para mejorar el rendimiento en una colección de 1 TB, simulando hasta 33 servidores. Los autores concluyen que el rendimiento de la replicación parcial con un broker de conexión mejora él de la aplicación de caching en el cliente o el servidor.

Finalmente, en [2] y [3] hemos estudiado el rendimiento de un sistema distribuido, replicado y basado en clusters simulando una colección Web como es SPIRIT [8]. Se identificaron dos principales cuellos de botella en un sistema distribuido y replicado: los brokers y la red. La carga en los brokers se debe al gran número de resultados parciales que deben ser ordenados. El problema en la red se debe al gran número de servidores de consulta y al continuo intercambio de datos con los brokers, especialmente en un sistema replicado. El análisis de un sistema basado en clusters indicaba que el mejor rendimiento de estos sistemas se obtiene cuando se utiliza un gran número de servidores de consulta, frente a un sistema replicado. Sin embargo, los sistemas basados en clusters deben ser configurados a-priori basándose en la distribución de las consultas que el sistema va a recibir.

3 Modelo de Simulación

Con el objetivo de estudiar el rendimiento de diferentes arquitecturas de un sistema distribuido de IR hemos implementado un simulador orientado a eventos discreto, utilizando JavaSim como entorno de simulación [10].

El modelo de simulación definido en este trabajo está dividido en tres partes. Inicialmente se describe un modelo analítico para la simulación de un sistema distribuido de IR, basado en nuestro trabajo previo en [2] y [3]. A continuación, se define un modelo para la colección de documentos con el objetivo de simular el comportamiento de cualquier colección, y en concreto una colección formada por 94 millones de documentos y 1 TB de texto. Finalmente, se presentan algunas extensiones a una red de área local básica.

3.1 Modelo Analítico

En un sistema distribuido de IR, las consultas se almacenan en una cola global controlada por uno o más *brokers centrales*. Cada broker tomará una consulta y la distribuirá a todos los servidores de consulta a través de la red, en un modelo de fichero invertido local [14]. Cada servidor de consulta deberá procesar localmente la consulta, obtener unos resultados locales para esa consulta, ordenador la lista de documentos, seleccionar un cierto número de los documentos más relevantes y enviárselos al broker. El broker recibirá todos los resultados parciales que serán combinados en una lista global y final de resultados.

En esta sección, describimos el modelo analítico para el proceso de búsqueda en un sistema distribuido de IR. El sistema simulado es una extensión del sistema básico descrito en [12].

Las variables básicas y los parámetros críticos del modelo analítico para un sistema distribuido de IR usando una estrategia de fichero invertido local, son los siguientes:

- q_i: vector de términos para la consulta *i-ésima*.
- k_i: número de términos en la consulta q_i.
- tc_1: primer coeficiente para el tiempo necesario para comparar e intercambiar dos identificadores.
- tc_2: segundo coeficiente para el tiempo necesario para comparar e intercambiar dos identificadores.
- ti: tiempo de inicialización, incluyendo reserva de memoria y salida de datos.
- ts: tiempo medio de posicionamiento para un disco.
- tr: tiempo medio para leer de disco la información sobre un documento en una lista invertida y su procesamiento (tiempo de posicionamiento no incluido).
- $d_{k,j}$: número de documentos en la lista invertida para el término k en el servidor de consulta j.
- $r_{i,j}$: número de resultados obtenidos para la consulta q_i en el servidor de consulta j.
- tr_{max}: número máximo de documentos devueltos como resultados parciales (se consideran únicamente los 1000 primeros documentos).
- $tr_{i,j}$: número de documentos del ranking en la consulta q_i devueltos como resultados parciales por el servidor de consulta j, donde $tr_{i,j} \le tr_{max}$.
- $t_{i,j}$: tiempo total (en milisegundos) para completar el procesamiento de la consulta q_i en el servidor de consulta j.
- $rq_{i,j}$: tiempo para recibir la consulta q_i por el servidor de consulta j.
- $ra_{i,j}$: tiempo para recibir los resultados parciales para la consulta q_i del servidor de consulta j.

Una vez que el servidor de consulta j recibe la consulta q_i para su procesamiento, lee de disco las listas invertidas asociadas con los términos k_i, cuya longitud viene dada por $d_{k,j}$. A continuación, las listas invertidas son combinadas y ordenadas para formar la lista de resultados parciales, cuya longitud viene dada por $r_{i,j}$. El tiempo necesario para combinar y ordenar n resultados, tc, se calcula como: $tc(n) = tc_1 \times n + tc_2 \times \ln(n)$. Por lo tanto, el tiempo necesario por el servidor de consulta j para procesar la consulta q_i viene dado por:

$$t_{i,j} = rq_{i,j} + ti + k_i \times ts + \sum_{k \in q_i} d_{k,j} \times tr + tc(r_{i,j})$$

Los valores usados en los parámetros t_i, t_s, t_r, tc_1, tc_2 fueron calculados a partir de un sistema real de IR y son, respectivamente, 1400, 0.03, 0.0040208, 0.00013068 y 0.000096 [2] [3]. Los parámetros $d_{k,j}$ y $r_{i,j}$ se estiman a través del modelo para la colección descrito en [2] y [3], para la colección SPIRIT.

Tan pronto como el broker ha recibido todos los resultados parciales de todos los servidores de consulta, debe combinarlos para obtener el resultado final. Por lo tanto, el tiempo total de procesamiento para la consulta q_i viene dado por:

$$t_i = \max(t_{i,j}) + \max(ra_{i,j}) + \sum_j tc(tr_{i,j}).$$

El problema es que los parámetros $rq_{i,j}$ y $ra_{i,j}$ no pueden ser estimados a través del modelo analítico, ya que dependen directamente de la carga de la red en cada momento. Por lo tanto, es necesario capturar el comportamiento de la red para representar de manera precisa los tiempos de respuesta de un sistema distribuido de IR.

3.2 Modelo de Red

En nuestros trabajos previos [2] y [3] el sistema distribuido de IR simulado se basaba en una única LAN que era representada por una cola FCFS de longitud infinita. Esta LAN gestionaba todos los mensajes enviados por los brokers a los servidores de consulta y viceversa.

Este modelo de red inicial presenta ciertas limitaciones que reducen las capacidades de los sistemas simulados. Este modelo de red representa una red de área local de acceso compartido, en donde todos los hosts de nuestro sistema distribuido de IR (servidores de consulta y brokers) están incluidos. Este sistema representa un entorno relativamente no realista, ya que no tiene en cuenta ciertas restricciones físicas de las LANs.

Con el objetivo de mejorar las limitaciones del modelo de red previo, hemos definido un nuevo modelo de red equivalente a una red conmutada Fast Ethernet 100BASE-T a 100 Mbps. Este nuevo modelo representa una red de área local conmutada en donde los hosts se interconectan a través de uno o más conmutadores (se asume que cada conmutador tiene capacidad para 64 hosts). También, se realiza una estimación de la sobrecarga exhaustiva, considerando las diferentes cabeceras de los protocolos de comunicación, fragmentación IP e

Tabla 1. Comparitiva entre los tiempos reales y estimados por el modelo de red

Tamaño mensajes	10 Mbps		100 Mbps	
	Mann-Whitney	Kolmogorov-Smirnov	Mann-Whitney	Kolmogorov-Smirnov
1000 bytes	0.889	1.000	0.898	1.000
2000 bytes	0.889	1.000	0.936	1.000
3000 bytes	0.894	1.000	0.997	1.000
4000 bytes	0.889	1.000	1.000	1.000

incluso el retardo de propagación. El diseño de este nuevo modelo de red también ha sido extendido para soportar tráfico multicast, frente al modelo previo que únicamente permitía mensajes unicast entre las diferentes máquinas.

De manera más detallada, el diseño del nuevo modelo de red está basado en una LAN conmutada. En este caso, cada host conectado a la red dispone de dos colas FCFS de longitud infinita (una para el envío y otra para la recepción de paquetes), comparado con el modelo previo basado en una única cola global. Este modelo permite representar una LAN conmutada, ya que dos hosts origen independientes pueden comunicarse simultáneamente con dos destinos independientes.

La cola de envío representa el buffer de salida de la interfaz de salida del transmisor, mientras que la cola de recepción representa el buffer de entrada de la interfaz del receptor. Para enviar un paquete, este se introduce en la cola de envío del host origen y se dirige directamente a la cola de recepción del destino.

La salida de un paquete desde la cola de envío a la cola de recepción se realiza cuando el paquete es el primer elemento de la cola y, si y sólo si, el receptor no está recibiendo otro paquete. En la cola de recepción, cada paquete tiene asignado un tiempo de servicio (antes de ser enviado al host) equivalente al tiempo de transmisión a través de la red, calculado como se describe a continuación.

Para la estimación del tiempo de transmisión de un paquete a través de la red se han considerado las cabeceras de los diferentes protocolos utilizados durante la comunicación y el retardo de propagación de la señal a través del sistema de transmisión.

El protocolo de comunicación utilizado por el sistema distribuido entre los servidores de consulta y los brokers se asume que está basado en un servicio no orientado a conexión (considerando el potencial número máquinas interconectadas y las características de los datos intercambiados), usando el protocolo UDP de la capa de transporte (con una cabecera de 8 bytes). El nivel de red utiliza IP (con una cabecera estándar de 20 bytes), y para el nivel de enlace se ha considerado el protocolo Ethernet (con una cabecera de 26 bytes).

Este modelo de red también considera la fragmentación IP de acuerdo al mecanismo descrito en el RFC 791 [13]. De manera resumida, si el tamaño del mensaje del nivel de red es mayor que el MTU (Maximum Transfer Unit) del nivel de enlace (1500 bytes en el caso de Ethernet), el mensaje original es dividido en fragmentos de un tamaño adecuado para el MTU. Estos fragmentos únicamente son reagrupados cuando alcanzan su destino final.

El retardo de propagación representa el tiempo que una señal electromagnética tarda en circular desde un extremo del segmento al otro. Este tiempo se mide normalmente en bits y depende en gran medida del tamaño del segmento y del método de transmisión utilizado. En una red FastEthernet 100BASE-T, el tiempo máximo para el retardo de propagación es de 512 bits (equivalente a 512 nanosegundos) [16]. En el modelo de simulación se ha escogido un valor intermedio de 256 bits (equivalente a 2.56 nanosegundos).

Finalmente, con el propósito de analizar la corrección del modelo propuesto para una LAN conmutada, se han desarrollado una serie de experimentos para comparar los tiempos de transmisión reales frente a los estimados mediante nuestro modelo.

En estos experimentos utilizamos dos ordenadores conectados a través de un cable cruzado, representando una conexión a través de un conmutador (el retardo introducido por un conmutador se considera despreciable). La conexión directa a través de un conmutador no fue posible, ya que los conmutadores disponibles no tenían buffers suficientes lo que provocaba la pérdida de paquetes.

Un proceso transmisor fue instalado en un ordenador y un proceso receptor en el otro. El transmisor era responsable de generar múltiples paquetes consecutivos (desde 10 a 1025), que eran recibidos por el proceso receptor responsable de la medida de los tiempos de respuesta.

En un experimento inicial, se evaluó la comunicación a través de una red Ethernet a 10 Mbps conectando dos Ultra Sparc 1 (128 MB RAM y un procesador a 167 MHz) y enviando mensajes de 1000, 2000, 3000 y 4000 bytes, midiendo los correspondientes tiempos

de envío. La equivalencia entre los tiempos reales y estimados fue evaluada utilizando los tests para dos muestras de Mann-Whitney y Kolmogorov-Smirnov, confirmando la equivalencia entre ambos con p-valores de más de 0.89 (ver Tabla 1 – izquierda).

También se evaluó una red FastEthernet a 100 Mbps, conectando dos PCs (512 MB RAM y un procesador a 1.5 GHz) y repitiendo los experimentos previos. La correspondencia entre los tiempos de transmisión reales y los estimados fue evaluada utilizando los tests de Mann-Whitney y Kolmogorov-Smirnov para dos muestras, confirmando la correspondencia entre ambos, con p-valores de más de 0.90 (ver Tabla 1 – derecha).

4 Resultados de la Simulación

Esta sección describe los resultados de varios experimentos desarrollados utilizando el modelo de simulación descrito en la sección anterior.

Todas las simulaciones están basadas en el modelo de colección SPIRIT [8]. Las consultas han sido modeladas utilizando el modelo de consultas sesgado [2][3] y siguiendo un escenario del peor caso posible: cada consulta recuperará, de media, 8.4 millones de documentos (el 9% del total de la colección). Un conjunto de 50 consultas es utilizado para medir el rendimiento, y para cada configuración se ejecutan 5 simulaciones diferentes (con distintas semillas iniciales), calculándose el valor medio de los tiempos de ejecución para cada consulta.

4.1 Sistema basado en Clusters

Un sistema basado en clusters está divido en grupos de ordenadores, en donde cada grupo opera como un sistema de IR autónomo. Cada cluster es responsable de una parte disjunta del total de la colección de documentos, y cada cluster puede utilizar distribución y replicación para almacenar su índice respectivo.

En [2] y [3], la principal conclusión demostraba que un sistema basado en clusters mejoraba el rendimiento de un sistema replicado, si se utilizaba un número elevado de servidores de consulta (p.e. 1024). Sin embargo, estos experimentos estaban basados en una red de acceso compartido, lo que provocaba la saturación de la red en un sistema replicado, y además, únicamente se consideraban cuatro réplicas en el sistema analizado.

Mediante estos experimentos, proporcionamos una comparativa detallada entre un sistema replicado y un sistema basado en clusters utilizando 1024 servidores de consulta y una red conmutada con soporte para multicast.

El sistema basado en clusters se define utilizando el trabajo de Spink et al. [15], en donde un conjunto de consultas reales de usuarios Web es clasificado en 11 temáticas diferentes, considerando tres años: 1997, 1999 y 2001. La Tabla 2 proporciona un resumen de las 11 temáticas y el porcentaje de consultas a través de los diferentes años.

Asumimos que cada temática está indexada en un cluster diferente. La colección se divide en 11 sub-colecciones con un fichero invertido de aproximadamente el mismo tamaño (8.5 millones de documentos) y por lo tanto los 11 clusters definidos indexarán el mismo número de documentos, aunque utilizando un número diferente de servidores de consulta. En estas simulaciones, el número de consultas ha sido incrementado a 200 y cada consulta recuperará de media 3 millones de documentos, para ajustarse al tamaño de los clusters.

La sub-colección base de 8.5 millones de documentos ha sido distribuida sobre N servidores de consulta utilizando una red conmutada y tres brokers, donde $N = 1, 2, 4, 8, 16, 32, 64, 128, 256$ y 512. En la Tabla 2, la columna *Configuración* describe los servidores de consulta asignados a cada temática. El primer número representa el número de servidores de consulta distribuidos y el segundo el número de réplicas en cada cluster.

El sistema basado en clusters ha sido configurado de acuerdo a la distribución temática del año 1997. El

Tabla 2. Distribución de las consultas en las diferentes temáticas y configuración del sistema basado en clusters simulado.

Temáticas	1997	1999	2001	Configuración
Entretenemiento	19.640 %	7.730 %	6.655 %	67 * 3
Pornografía	16.540 %	7.730 %	8.555 %	56 * 3
Comercio	13.030 %	24.730 %	24.755 %	66 * 2
Informática	12.240 %	11.130 %	9.654 %	63 * 2
Ciencias	9.240 %	8.020 %	7.554 %	48 * 2
Gente	6.430 %	20.530 %	19.754 %	66 * 1
Sociedad	5.440 %	4.430 %	3.955 %	56 * 1
Educación	5.330 %	5.520 %	4.554 %	55 * 1
Arte	5.140 %	1.330 %	1.155 %	53 * 1
No en Inglés	3.840 %	7.030 %	11.355 %	39 * 1
Gobierno	3.130 %	1.820 %	2.054 %	32 * 1

Tabla 3. Throughput (consultas/segundo) y tiempo de respuesta (milisegundos) para el sistema basado en clusters y los sistemas replicados.

Replicado	Throughput	Tiempo de Respuesta	*Cluster* *Brokers*	*Año*	Throughput	Tiempo de Respuesta
1x1024	0.70	4247.83	*54*	*1997*	7.60	2404.11
2x512	1.38	4257.67	*54*	*1999*	3.23	2828.11
4x256	2.69	3231.22	*54*	*2001*	3.59	2960.87
8x128	5.03	2354.92	*37*	*1997*	7.17	2380.20
16x64	8.47	2274.09	*37*	*1999*	3.11	3165.59
32x32	12.92	2658.93	*37*	*2001*	3.43	2863.65

número de réplicas para los temas más populares ha sido maximizado, manteniendo el número de servidores de consulta en cada réplica lo más cerca posible a 64 para obtener tiempos de respuesta adecuados. El número de brokers utilizados ha sido seleccionado considerando el total de réplicas en cada cluster (p.e. *R=18* réplicas) y calculando el número óptimo de brokers obtenido en este trabajo (*3R*) y en nuestro trabajo previo (*2R+1*).

El rendimiento de este sistema basado en clusters es comparado con un sistema replicado, estudiando diferentes configuraciones: *1, 2, 4, 8, 16 y 32* réplicas (con 1024, 512, 256, 128, 64 y 32 servidores de consulta por réplica, respectivamente), siempre utilizando el número óptimo de brokers.

En estos experimentos, se mide el rendimiento utilizando el throughput y el tiempo de respuesta. El throughput se mide considerando que el sistema dispone de una cola de 200 consultas, que serán procesadas de manera secuencial y consecutiva (modo batch). El tiempo de respuesta se mide considerando que las consultas son recibidas por el sistema de IR según una distribución Exponencial, con media 500 milisegundos y simulando 200 consultas.

Los resultados obtenidos para el sistema basado en clusters y los sistemas replicados se presentan en la Tabla 3.

Con respecto a los sistemas replicados, observamos que el throughput aumenta según aumenta el número de réplicas debido al mayor nivel de paralelismo del sistema. Simultáneamente, el tiempo de respuesta del sistema decrece según aumenta el nivel de replicación, excepto en la última configuración en donde la reducida distribución del índice (únicamente se utilizan 32 servidores de consulta por réplica) incrementa el tiempo de respuesta.

Por otra parte, el throughput de un sistema basado en clusters se maximiza cuando se utilizan *3R* brokers. Otro punto importante es la significativa reducción en el rendimiento que se produce ante un cambio de tendencia en las consultas (el throughput se reduce en más del 50% y el tiempos de respuesta se incrementan aproximadamente un 17%). Esto es debido a los importantes cambios en la distribución de las temáticas desde el año 1997 (base para los actuales experimentos) hasta los años 1999 y 2001. En [2] y [3] se tomaba como base la distribución para el año 2001, y la reducción en el rendimiento se ponía de manifiesto con la distribución del año 1997, con cambios mínimos en el año 1999.

Comparando los dos tipos de sistemas, los resultados demuestran que un sistema replicado con 16 réplicas obtendrá un mejor throughput y tiempo de respuesta que el sistema basado en clusters definido. En ambos casos el nivel de paralelismo es similar (16 y 18 réplicas, respectivamente).

El principal beneficio del sistema basado en clusters es una reducción en el tráfico de la red, que es crucial si la red es el principal cuello de botella del sistema. Sin embargo, la utilización de una red conmutada ha resuelto este problema, optimizando significativamente el rendimiento de los sistemas replicados, lo que les permite mejorar el rendimiento de los sistemas basados en clusters.

5 Conclusiones

Este artículo es la continuación de nuestro estudio previo sobre diferentes arquitecturas para sistemas de IR distribuidos, presentado en [2] y extendido en [3]. En este trabajo hemos proporcionado un estudio detallado de un sistema basado en clusters comparando su rendimiento (midiendo throughput y tiempo de respuesta) con varios sistemas replicados.

Basándonos en los experimentos de la red, la utilización de una red conmutada para un sistema distribuido proporciona mejoras en el rendimiento en todos los casos. Sin embargo, esta mejora es más significativa en los sistemas replicados, ya que en estos casos el cuello de botella lo constituye la red de interconexión. La red conmutada evitará la saturación de la red lo que permitirá mejorar considerablemente el rendimiento, especialmente en comparación con una red de acceso compartido saturada con un gran número de servidores de consulta.

Otras alternativas, como la utilización de mensajes multicast o la utilización de conexiones a 1 Gbps con los brokers proporcionan mínimas mejoras en el

rendimiento, ya que la red ha dejado de ser el principal cuello de botella del sistema replicado, pasando a ser los brokers.

En lo que respecta al sistema basado en clusters analizado, su rendimiento (medido tanto en throughput como en tiempo de respuesta) no mejora los valores obtenidos por el mejor sistema replicado. Este resultado, que en cierta medida contradicen las conclusiones de nuestro trabajo previo, está relacionado con el uso de una red conmutada.

La principal ventaja de un sistema basado en clusters es la reducción del tráfico en la red. Sin embargo, la red conmutada ha eliminado el cuello de botella de la red, mejorando significativamente el rendimiento de los sistemas replicados, que supera el obtenido por lo sistemas basados en clusters. Sin embargo, es importante destacar que hay otros factores que no han sido tenidos en cuenta (como la reducción en el número de documentos relevantes producida en los clusters) que pueden mejorar el rendimiento de los sistemas basados en clusters.

Sin embargo, el objetivo final de este trabajo, al igual que en nuestro trabajo previo, es poder utilizar estos resultados para convertir un sistema de IR centralizado en uno distribuido, y analizar la correspondencia entre el rendimiento estimado y el real.

De manera genérica, consideramos que los resultados de este artículo pueden ser útiles para cualquier grupo interesado en gestionar una gran colección como SPIRIT o para construir un motor de búsqueda a gran escala.

Agradecimientos

Quisiéramos agradecer la inestimable ayuda prestada por Alejandra Barreiro y Alejandro Moratinos en el desarrollo de los experimentos de red desarrollados.

Referencias

[1] Burkowski, F. J. (1990). Retrieval performance of a distributed database utilising a parallel process document server. In *Proceedings of the second International Symposium on Databases in Parallel and Distributed Systems* (pp. 71-79). New York: ACM Press.

[2] Cacheda, F., Plachouras & V., Ounis, I. (2003). Performance Analysis of Distributed Architectures to Index One Terabyte of Text. In *Proceedings of 26th European Conference on Information Retrieval Research (ECIR'04)*, Lecture Notes on Computer Science (2997), pp. 394-408.

[3] Cacheda, F., Plachouras, V. & Ounis, I. (2005). A Case Study of Distributed Information Retrieval Architectures to Index One Terabyte of Text. *Information Processing and Management Journal*, Volume 41, Issue 5, pp: 1141-1161, 2005.

[4] Cahoon, B. & McKinley, K.S. (1996). Performance evaluation of a distributed architecture for information retrieval. In *Proceedings of 19th ACM-SIGIR International Conference on Research and Development in Information Retrieval* (pp: 110-118). New York: ACM Press.

[5] Harman, D., Mccoy, W., Toense, R. & Candela, C. (1997). Prototyping a distributed information retrieval system that uses statistical ranking. *Information Processing and Management*, 27 (5), 449-460.

[6] Hawking, D. (1997). Scalable text retrieval for large digital libraries. *Lecture Notes in Computer Science*, 1324, 127-146.

[7] Hawking, D. & Thistlewaite, P. (1999). Methods for Information Server Selection. *ACM Transactions on Information Systems*, 17(1), 40-76.

[8] Jones, C. B., Purves, R., Ruas, A., Sanderson, M., Sester, M., van Kreveld, M. & Weibel, R. (2002). Spatial information retrieval and geographical ontologies an overview of the SPIRIT project. In *Proceedings of the 25th ACM-SIGIR Conference on Research and Development in Information Retrieval*, (pp. 387-388). New York: ACM Press.

[9] Lin, Z. & Zhou, S. (1993). Parallelizing I/O intensive applications for a workstation cluster: a case study. *ACM SIGARCH Computer Architecture News*, 21 (5), 15-22.

[10] Little, M. C. (2001). *JavaSim User's Guide. Public Release 0.3, Version 1.0*. University of Newcastle upon Tyne. Retrieved 1 June, 2003, from the World Wide Web: http://javasim.ncl.ac.uk/manual/javasim.pdf

[11] Lu, Z. & McKinley, K. (2000). Partial collection replication versus caching for information retrieval systems. In *Proceedings of the 25th ACM-SIGIR Conference on Research and Development in Information Retrieval*, (pp. 248-255). New York: ACM Press.

[12] Plachouras, V., Ounis, I., Amati, G. & van Rijsbergen C.J. (2002). University of Glasgow at the Web track of TREC 2002. In Voorhees E.M. & Buckland, L.P. (Ed.), *Proceedings of The Eleventh Text REtrieval Conference* (pp. 645-651). Gaithersburg, Maryland: NIST Special Publication 500-251.

[13] Postel, J. (1981). Internet Protocol. RFC 791, Internet Engineering Task Force.

[14] Ribeiro-Neto, B. & Barbosa, R. (1998). Query performance for tightly coupled distributed digital libraries. In *Proceedings of the 3rd ACM Conference on Digital Libraries*, (pp: 182-190). New York: ACM Press.

[15] Spink, A., Jansen, B. J., Wolfram, D. & Saracevic, T. (2002). From e-sex to e-commerce: Web search changes. *IEEE Computer* 35(3): 107-109.

[16] Spurgeon, C.E. (2000). Ethernet: The Definitive Guide. Sebastopol: O'Reilly & Associates.

[17] Tomasic, A. & Garcia-Molina, H. (1993). Performance of inverted indices in shared-nothing distributed text document information retrieval systems. In *Proceedings of the 2nd International Conference on Parallel and Distributed Information Systems*, (pp: 8-17). San Diego, California: IEEE Computer Society.

Definición y desarrollo de un sistema de generación de reglas XPath dinámico para la creación de extractores de documentos HTML de la Web

Fernando Paniagua Martín, Vicente Luque Centeno
Departamento de Ingeniería Telemática. Universidad Carlos III de Madrid
Avenida de la Universidad, 30.
28911 – Leganés (Madrid)
E-mail: fernando.paniagua@gmail.com, vlc@it.uc3m.es

Abstract. *The automatic recovery of information from the Internet can be done by information extractors based on XPath rules. These extractors must work over HTML pages previously converted to XHTML. The creation of XPath rules can be quite complex and requires to have very specific knowledge. On the other hand and due to the size of the documents we are working with, the duty could be very laborious. In this work, a solution to the XPath rules creation by the definition and development of a graphical tool for the automatic and configurable generation of these rules has been implemented. This tool covers the rule generation process too. This process goes from the conversion of the HTML documents to XHTML, to the verification of this conversion.*

1 Introducción

La World Wide Web se concibe inicialmente como una tecnología para compartir información entre personas. La principal motivación que tienen los primeros usuarios de Internet es publicar información que sea accesible para otros usuarios. Pero Internet crece, se hace prácticamente universal y la cantidad de datos que se publican se hace inmensa. Esto se traduce en la necesidad de poder procesar de manera automática la información disponible debido a que su volumen la hace inmanejable.

La gran mayoría de contenidos publicados en Internet están expresados en lenguaje HTML y esto supone algunos inconvenientes. Una página HTML no proporciona, por definición, información semántica sobre su contenido. Las etiquetas del lenguaje HTML sólo proporcionan información sobre la presentación y no sobre qué significan los datos que contienen. Esta limitación se subsanaría fácilmente con la utilización de XML como lenguaje para la publicación de contenidos siempre y cuando se aprovechase para dotar a las diferentes partes de cada documento de información semántica. Si la publicación de contenidos se hiciese de esta manera el procesamiento de las páginas publicadas sería relativamente sencillo al poder identificar de manera inequívoca el significado de cada dato. Esto no es así en la gran mayoría de los casos. Esto probablemente se deba al desconocimiento de las tecnologías y a las ventajas que se pueden obtener de publicar la información de manera estructurada. En muchos casos resultará interesante a los autores o propietarios de determinada información que esta participe en los procesos de extracción automáticos. En algunos otros casos, en cambio, es probable que la información nunca sea publicada de tal manera voluntariamente para evitar dar facilidades a los procesos de extracción automáticos ya que, en ocasiones, la financiación de los sitios Web está íntimamente ligada al número de visitas reales de sus páginas.

Para realizar el procesamiento de una página HTML hay, por lo tanto, que fijarse en la información que envuelve a los datos: aquella referente a la presentación. Es razonable pensar que un autor seleccionará un formato de presentación llamativo para un dato que tiene gran relevancia mientras que otro dato menos importante tendrá un formato que llame menos la atención. Al fin y al cabo se está mostrando información que será consultada por personas y los autores de las páginas han de decidir sobre qué prefieren que los destinatarios de la información fijen su atención.

Dadas estas premisas se puede afirmar que a partir del conocimiento de la información relacionada con la presentación de los datos se podrían realizar procesos de extracción de los mismos para su tratamiento posterior. Este tipo de procesos se podría realizar de diversas formas. En este caso se ha optado por trabajar con documentos XML generados a partir de las páginas HTML para poder de esta forma utilizar estándares abiertos y librerías freeware que se basan o trabajan con XML: Lenguaje de programación Java, la librería JTidy como herramienta de conversión de documentos HTML en documentos XML, XPath como lenguaje para recuperar información de un documento XML o DOM como procesador de documentos XML.

Para hacer efectiva esta solución es necesario resolver tres problemas:

1. Un documento HTML no necesariamente cumple las rígidas normas de XML y, por lo tanto, no se contempla la posibilidad de poder realizar consultas

con reglas XPath sobre él. Además los navegadores más utilizados por los usuarios de Internet se han desarrollado de tal forma que 'relajan' bastante las exigencias sobre los documentos HTML. Esto ha provocado que un gran número de documentos HTML no cumplan con su gramática. Todo esto hace necesario disponer de un mecanismo capaz de convertir una página HTML en un documento XML equivalente.

2. Los documentos HTML que se pueden encontrar en Internet pueden tener un tamaño enorme. El procesamiento manual del documento XML resultante de la conversión puede ser bastante complicado y engorroso.

3. XPath es un lenguaje para seleccionar partes de un documento XML que no es demasiado sencillo de utilizar para un usuario no experimentado.

Es posible realizar un proceso de extracción de una página HTML basándose únicamente en el procesamiento de cadenas de caracteres. Una página HTML es, sencillamente, una gran cadena de caracteres y un algoritmo realizado con cierta habilidad permitiría extraer información de forma automatizada. La dificultad surge cuando tratamos distintas páginas totalmente heterogéneas en cuanto a estructura y contenido.

No es recomendable utilizar un algoritmo distinto para procesar cada una de las páginas porque tiene como desventajas que requerirá un gran esfuerzo de programación así como que impedirá a personas sin los debidos conocimientos en programación la adaptación del proceso a sus necesidades.

Por otra parte hay que intentar que la flexibilidad del proceso de generación de reglas no resida en el algoritmo sino en la configuración del mismo. De esta forma será posible cambiar el comportamiento del proceso de generación sin que sea necesario modificar la aplicación.

Utilizando XPath se va a trasladar la robustez del proceso a las reglas: cambiando la regla se cambia el comportamiento del proceso y eso habilita a un amplio abanico de usuarios con conocimientos menos específicos a modificar el proceso según sus necesidades. Lamentablemente, y como se comentaba en el punto 3, la elección de las reglas es un proceso complejo y proporcionar ayuda en este punto es, concretamente, una de las principales motivaciones de este trabajo.

La generación de estas reglas XPath es, en resumen, el proceso más complicado pero a la vez el más importante: el éxito en la elección de las reglas que seleccionan la parte del documento que deseamos para su procesamiento posterior determinará el éxito del proceso de extracción basado en dichas reglas.

2 Herramientas de visualización de documentos XML

Existe un gran número de herramientas y utilidades para el manejo de documentos XML. La funcionalidad que proporcionan es diversa si bien ninguna de dichas herramientas es capaz de generar conjuntos de reglas XPath de manera dinámica. Algunas de ellas son capaces de generar reglas XPath pero estas suelen ser de poca utilidad para la creación de extractores ya que son extremadamente precisas y suelen servir únicamente para localizar el nodo seleccionado. A continuación se presentan algunas de estas soluciones para la representación de documentos XML y, en algunos casos, la creación y/o la ejecución de reglas XPath sobre los mismos.

- Algunos de los navegadores más utilizados como Mozilla Firefox, Netscape o Internet Explorer representan documentos XML y realizan la transformación con XSL.

- “Exchange XML Browser”[1] es un visor de documentos XML. Permite introducir y ejecutar reglas XPath.

- Amaya [2]. Es un editor Web que permite visualizar y editar documentos XML.

- Jumbo [3]. Es un visor de documentos escritos en CML (Chemical Markup Language) y en XML.

- X-Smiles [4]. Es un visor de documentos XML para cualquier tipo de dispositivo que sea capaz de ejecutar aplicaciones Java. Uno de sus objetivos es realizar transformaciones XSL.

- XML Converter [5]. Esta herramienta permite al usuario crear de manera interactiva una transformación de datos.

- XML Fox [6]. Es una herramienta gráfica para la creación de documentos XML.

- XpathVisualizer [7]. Es un visor de documentos XML que permite la ejecución de reglas XPath. Proporciona una ayuda para la generación de las reglas.

- XSLDebugger [8]. Permite la depuración visual de transformaciones XSL.

- XPath Query Expression Tool [9]. Página HTML que permite la ejecución de una sentencia XPath sobre un documento XML. La representación del documento XML no es visual y no proporciona ninguna ayuda.

- XPath Explorer [10] permite la ejecución de una expresión XPath sobre un documento XML o HTML. También incluye la posibilidad de generar una expresión XPath capaz de devolver un nodo. La expresión generada no es genérica y sólo devolverá el nodo seleccionado. No permite la búsqueda de nodos a partir de un texto.

- Stylus Studio [11] permite buscar nodos a partir de un texto, es capaz de trabajar con documentos XML y HTML, pero sólo ofrece una única expresión XPath como resultado de la selección de un nodo. La regla generada es absoluta y únicamente devolverá el nodo seleccionado.

3 Sistema de Generación de Reglas XPath

El proceso de creación de extractores de información basados en reglas XPath se puede apoyar en algunas de las aplicaciones y herramientas mencionadas anteriormente. Dichas herramientas permiten la ejecución de manera visual de reglas XPath lo que permite comprobar que los resultados esperados son los correctos. Pero no ayudan al problema de la creación de las reglas, lo que podría resultar de lo más útil para hacer accesible esta tecnología a un mayor número de usuarios.

Por otra parte hay que destacar que las herramientas disponibles no cubren más que algunas de las posibles fases por las que habría que transitar a la hora de construir un extractor. Esto supone una dificultad añadida que se trata de resolver con el sistema desarrollado.

En este trabajo se han identificado los pasos de los que consta el proceso de preparación del juego de reglas XPath necesario para la creación de un extractor de información de la Web. En la figura 1 se puede observar las tareas que componen el proceso.

El proceso se inicia en un documento, normalmente escrito en lenguaje HTML, sobre el que se desea encontrar un conjunto de reglas XPath que sean capaces de recuperar la información deseada. HTML es el lenguaje en el que se escriben la mayor parte de los documentos publicados en la Web.

Lamentablemente la ejecución de reglas XPath sobre documentos HTML no es posible ya que XPath sólo trabaja sobre documentos XML. Esto obliga a realizar una transformación del documento HTML original en un documento XML equivalente. En concreto los documentos HTML se convierten a XHTML que es un tipo de documento XML. Esta conversión será, por lo tanto, el primer paso a realizar.

Una vez se dispone de un documento XHTML se procede a su visualización. Un documento XML es un árbol y esa es la manera de representarlo más

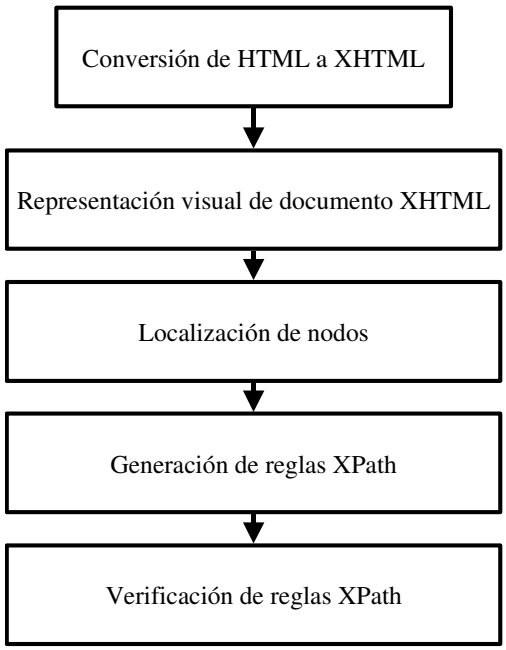


Fig. 1. Pasos a seguir para la creación y prueba de reglas XPath

intuitiva y sencilla. Se pretende con la representación visual dotar al sistema de un mayor grado de interactividad y sencillez.

El tercer paso consiste en la localización de los nodos de referencia. Cuando se quiera generar una regla XPath para que devuelva un determinado tipo de información hay que, en primer lugar, indicar qué nodo será tomado como ejemplo y referencia para realizar la generación.

A partir del nodo de referencia se generan un conjunto de reglas XPath. Todas estas reglas aplicadas sobre el documento XHTML cargado en el sistema devolverán, al menos, el nodo seleccionado. Por similitud algunas de estas reglas devolverán el nodo de referencia y aquellos que tengan una apariencia similar. El conjunto de reglas devueltas es variable e irá en función de la configuración del sistema.

Por último se permite la validación de las reglas en el propio sistema. Una vez generadas se podrán ejecutar sobre el documento cargado para comprobar si los resultados son o no los esperados.

3.1 Conversión de HTML a XHTML

Esta conversión ya estaría soportada por algunas librerías existentes como Tidy [12] o su versión Java JTidy [13]. En este trabajo se ha integrado la utilización de JTidy para que no se requiera convertir de manera externa el documento HTML con el que se va a trabajar.

La conversión con JTidy no es inmediata. JTidy no es capaz de convertir a XHTML todos los documentos HTML disponibles en la Web. Algunas páginas HTML contienen información que JTidy no sabe cómo resolver y hace que el proceso de conversión

termine de forma errónea. Ha sido necesario realizar un proceso previo en el que se preparan las páginas eliminando la información que se sabe a priori que provocará error en el proceso de conversión.

3.2. Representación visual del documento XML

Como se ha detallado en la introducción existen multitud de visores de documentos XML. Se ha modificado un visor de documentos XML disponible en la Web de Java de Sun [14] para añadir la nueva funcionalidad.

La representación visual es una necesidad para poder trabajar de manera rápida con este tipo de documentos. La alternativa sería trabajar con el código del documento y, debido a su tamaño, es poco práctica.

En la figura 2 se muestra una captura de la pantalla principal del sistema en el que se puede observar un documento XHTML cargado en el sistema. Al tratarse de un documento XML la manera de representarlo más intuitiva es en forma de árbol.

Por otra parte hay que indicar que el sistema permite trabajar con ficheros recuperados a partir de una URL o del propio sistema de ficheros local.

3.3 Localización de nodos

Una aportación de este trabajo es añadir un buscador de nodos que contengan un determinado texto. Partiendo de la premisa de que los extractores se utilizan para la recuperación de textos parece razonable que la tarea de creación de reglas XPath se base en los mismos.

La localización de los nodos se realizará de la siguiente manera: desde el navegador de Internet el usuario identificará en la página HTML original un texto representativo del tipo de información que desee recuperar posteriormente (un titular de una noticia, por ejemplo). Este texto representativo se introducirá en el sistema y este buscará todos aquellos nodos que lo contienen. A continuación el sistema mostrará una lista con todos los nodos resultantes ya que puede ser que el texto elegido se encuentre en más de un punto en el documento. El usuario podrá seleccionarlos de uno en uno comprobando cuál se corresponde con el nodo elegido.

En la figura 3 se muestra cómo la búsqueda de los nodos que contienen una determinada cadena de caracteres devuelve dos elementos como respuesta. Es responsabilidad del usuario decidir cuál de ellas es la correcta o elegir un texto diferente que genere únicamente la respuesta deseada.

Esta utilidad resulta especialmente interesante debido a que el tamaño de los documentos HTML puede hacer que la localización manual de los nodos de referencia resulte compleja. La alternativa a la localización de nodos de manera automática consistiría en buscar la cadena de caracteres deseada dentro del código del documento para, a continuación, extender el árbol completamente hasta localizar el nodo que la contiene. Como ya se ha indicado anteriormente el tamaño de los documentos HTML con los que normalmente se va a trabajar es

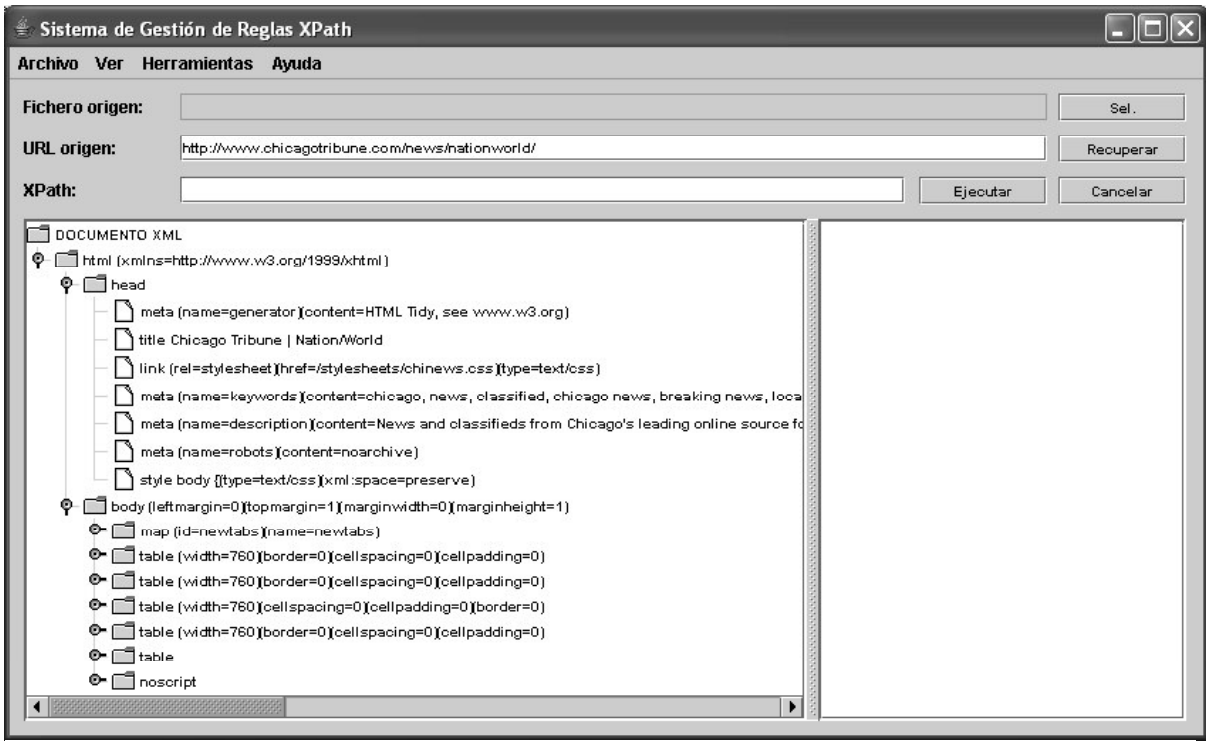


Fig 2: Pantalla principal del sistema

tal que hacen que la localización manual de los nodos sea difícil de realizar.

3.4. Generación de reglas XPath

La generación de expresiones se realiza a partir de la selección de un nodo en el visor de documentos XML.

Con la información del nodo seleccionado y con la obtenida en el fichero de configuración se lleva a cabo la generación de expresiones. Por cada condición (cada nodo *candidate* del fichero de configuración) que cumpla el nodo seleccionado se generará una expresión.

En el fichero de configuración se indica qué tipo de expresiones se pueden generar. Existen tres posibilidades:

- Para indicar al sistema que debe generar expresiones basándose en la existencia de determinado atributo se ha de proporcionar la siguiente información:

  - *Atributo de referencia.*
  - *Número de nodos a examinar.*

  Ejemplo:

  <candidate type="attribute" name="class" number_of_nodes="3" />

  Lo que se está diciendo al sistema es que tome como expresión candidata aquella que lleve una secuencia de tres nodos (el seleccionado y los dos que antecesores) si en uno de ellos (indistintamente) se encuentra el atributo *class*. La expresión generada tendrá un formato similar al siguiente, pudiendo variar de posición la referencia al atributo:

  //nodo1/nodo2[class='valor']/nodo_elegido

- Por otro lado si se desea generar expresiones en las que se tome como referencia un nodo determinado y un número de nodos antecesores se deberá proporcionar la siguiente información:

  - *Nodo de referencia.*
  - *Número de nodos a seleccionar.*

  Ejemplo:

  <candidate type="simple_node" name="td" number_of_nodes="3" />

  En este caso, lo que se está notificando al sistema es que genere una expresión XPath candidata si el nodo seleccionado es del tipo *td* en la que se incluyan el nodo antecesor y el antecesor de este último. La expresión generada tendría el formato siguiente:

  //antecesor_de_antecesor/antecesor/td

- Como última alternativa implementada se puede indicar a la aplicación que genere expresiones XPath tomando nodos de manera directa, sin validar ni tener en cuenta ninguna información

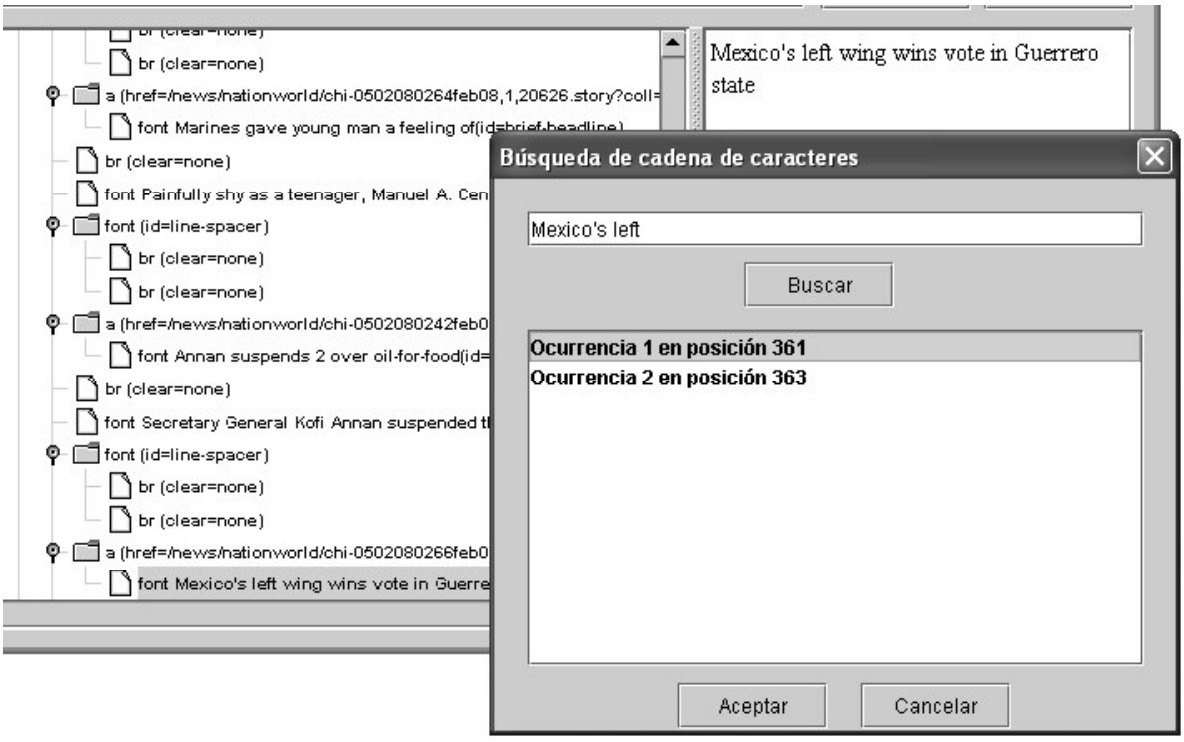


Fig. 3. Ejemplo de localización de nodos

salvo el número de nodos a tomar para formar la expresión. La información que se debe proporcionar en este caso es la siguiente:

- *Número de nodos a seleccionar.*

Ejemplo:

<candidate type="simple_node" name="" number_of_nodes="3"/>

Con este dato la expresión XPath candidata incluiría el nodo seleccionado, el antecesor y el antecesor de este último. La expresión generada tendría la apariencia que se puede ver a continuación:

//antecesor_de_antecesor/antecesor/nodo_seleccionado

De esta forma el número de expresiones que se generarán a partir de la selección de un nodo dependerá de las condiciones indicadas en el fichero de configuración así como de las características del propio nodo. El fichero de configuración es un documento XML que deberá ser modificado cuando el conjunto de expresiones obtenidas no sea suficiente por no generar reglas lo suficientemente exhaustivas y precisas.

Las expresiones generadas de esta manera, al ser aplicadas sobre un documento XML devolverán un conjunto de nodos que coincidirán con aquellas características del nodo de referencia que cumplían con las condiciones indicadas en la configuración. Estas expresiones no incluyen el total de expresiones que devolverían el mismo conjunto de nodos debido a que el proceso de generación ha sido desarrollado dentro de unos límites. En la mayor parte de las pruebas realizadas el algoritmo creado ha demostrado ser lo suficientemente potente cumpliendo con las necesidades que existían en cada caso. Además otro tipo de expresiones como la que iría desde la raíz hasta el nodo seleccionado han de ser descartadas ya que son muy precisas pero muy poco exhaustivas.

El fichero de configuración es un documento XML que deberá cumplir con el siguiente DTD, en el cual se contemplan las tres posibilidades anteriormente:

```
<!ELEMENT candidates (candidate)+>
<!ELEMENT candidate EMPTY>
<!ATTLIST candidate type (attribute|simple_node) #REQUIRED>
<!ATTLIST candidate name CDATA #REQUIRED>
<!ATTLIST candidate number_of_nodes CDATA #REQUIRED>
```

La alternativa a la generación automática de reglas consiste en analizar un nodo de referencia y sus nodos antecesores de manera manual. Una vez hecho esto se deberá escribir una regla suficientemente genérica de tal manera que su ejecución devuelva los nodos que contienen la información deseada. La creación de las reglas es un trabajo que lleva tiempo realizar y que, además, requiere de conocimientos específicos de HTML y de XPath lo que hace que el número de personas que puedan realizarlo esté limitado.

Según esta información se puede ver como partiendo de la siguiente configuración del proceso de generación de reglas se pueden obtener un conjunto útil de reglas XPath:

```
<candidates>
<candidate type="attribute" name="shape" number_of_nodes="1"/>
<candidate type="attribute" name="shape" number_of_nodes="2"/>
<candidate type="attribute" name="class" number_of_nodes="0"/>
<candidate type="attribute" name="class" number_of_nodes="1"/>
<candidate type="attribute" name="class" number_of_nodes="2"/>
<candidate type="attribute" name="class" number_of_nodes="3"/>
<candidate type="attribute" name="bgcolor" number_of_nodes="3"/>
<candidate type="simple_node" name="" number_of_nodes="0"/>
<candidate type="simple_node" name="" number_of_nodes="1"/>
<candidate type="simple_node" name="" number_of_nodes="2"/>
</candidates>
```

En el siguiente código se puede observar el fragmento de página HTML en el que se encuentra resaltado en negrita el texto del nodo que se ha seleccionado para la generación de las reglas.

```
<tr valign="top">
<td width="339">
<div class="diezyochopixazul">
<a href="/11568.html" class="diezyochopixazul" shape="rect">
TITULAR
</a>
</div>
</td>
</tr>
```

El conjunto de reglas obtenido como resultado es el siguiente:

```
//div/a[@shape='rect']
//td/div/a[@shape='rect']
//a[@class='diezyochopixazul']
//div[@class='diezyochopixazul']/a[@class='diezyochopixazul']
//td/div[@class='diezyochopixazul']/a[@class='diezyochopixazul']
//tr/td/div[@class='diezyochopixazul']/a[@class='diezyochopixazul']
//a
//div/a
//td/div/a
```

La relación entre las entradas del fichero de configuración y las reglas generadas se muestra a continuación en bloques por tipo de condición de generación:

Entradas:

```
<candidate type="attribute" name="shape" number_of_nodes="1"/>
<candidate type="attribute" name="shape" number_of_nodes="2"/>
```

Reglas generadas:

```
//div/a[@shape='rect']
//td/div/a[@shape='rect']
```

Entradas:

```
<candidate type="attribute" name="class" number_of_nodes="0"/>
<candidate type="attribute" name="class" number_of_nodes="1"/>
<candidate type="attribute" name="class" number_of_nodes="2"/>
<candidate type="attribute" name="class" number_of_nodes="3"/>
```

Reglas generadas:

```
//a[@class='diezyochopixazul']
//div[@class='diezyochopixazul']/a[@class='diezyochopixazul']
//td/div[@class='diezyochopixazul']/a[@class='diezyochopixazul']
//tr/td/div[@class='diezyochopixazul']/a[@class='diezyochopixazul']
```

Entradas:

```
<candidate type="attribute" name="bgcolor" number_of_nodes="3"/>
```

Reglas generadas:

```
No genera reglas.
```

Entradas:

```
<candidate type="simple_node" name="" number_of_nodes="0"/>
<candidate type="simple_node" name="" number_of_nodes="1"/>
<candidate type="simple_node" name="" number_of_nodes="2"/>
```

Reglas generadas:

```
//a
//div/a
//td/div/a
```

Como se puede ver en el desglose anterior no todas las entradas del fichero de configuración generarán reglas. Las entradas son condiciones y en función del nodo seleccionado estas condiciones darán paso a nuevas reglas o no.

En la figura 4 se muestra un ejemplo de generación de reglas en la pantalla principal del sistema a partir de la selección de un nodo.

3.5 Verificación de reglas XPath

Por último se ha dotado al sistema de la posibilidad de verificar las reglas XPath sobre el documento XML cargado en el visor. Introduciendo una regla XPath manualmente o seleccionándola de entre las generadas se podrá ejecutar sobre el documento con el que se está trabajando. A continuación se mostrarán en el árbol los nodos resultantes de la ejecución. De esta forma se podrá comprobar, de manera visual, que los resultados obtenidos de las reglas son los esperados y que, por lo tanto, pueden pasar a formar parte de un futuro extractor.

En la figura 5 se muestra como se presenta el resultado de la ejecución de una regla XPath sobre un documento. El árbol con el documento original es sustituido por un árbol en el que aparecen los nodos resultantes de la ejecución de la regla. En este punto el usuario deberá comprobar si estos nodos se corresponden con los resultados esperados para determinar si la regla es válida y las probabilidades de que tenga éxito al ser utilizada en la creación de un extractor.

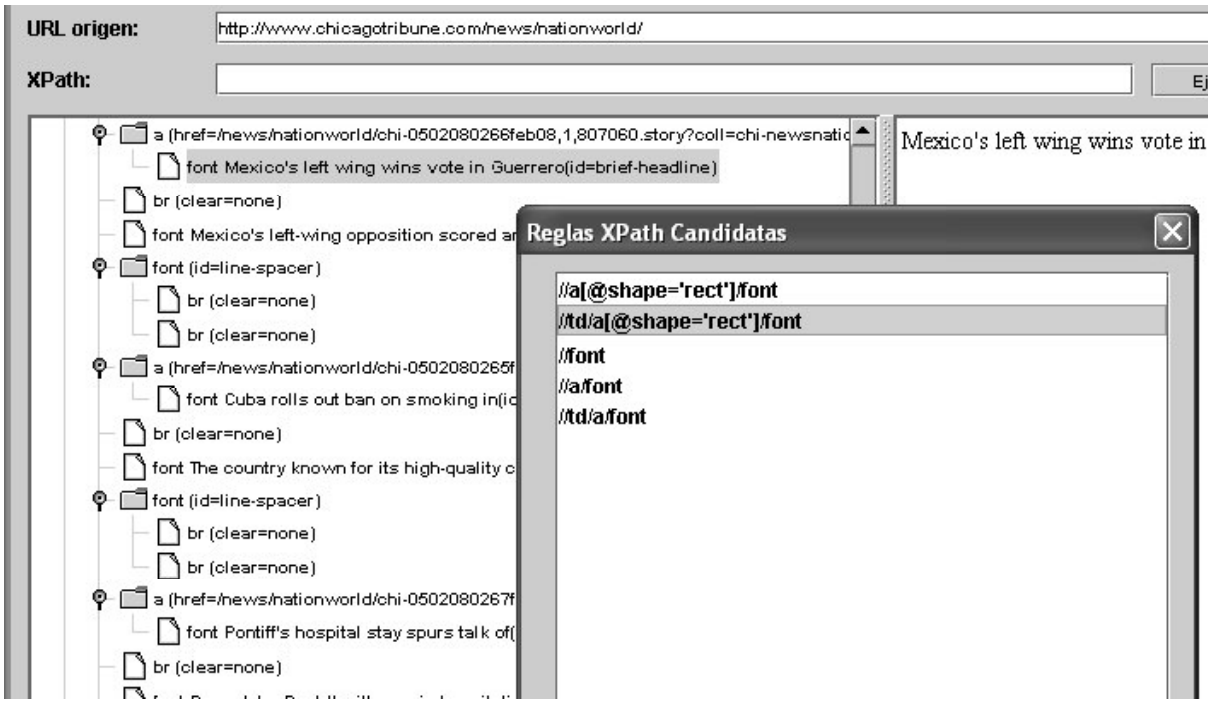


Fig. 4. Visor de documentos XML y ejemplo de reglas XPath generadas por el sistema

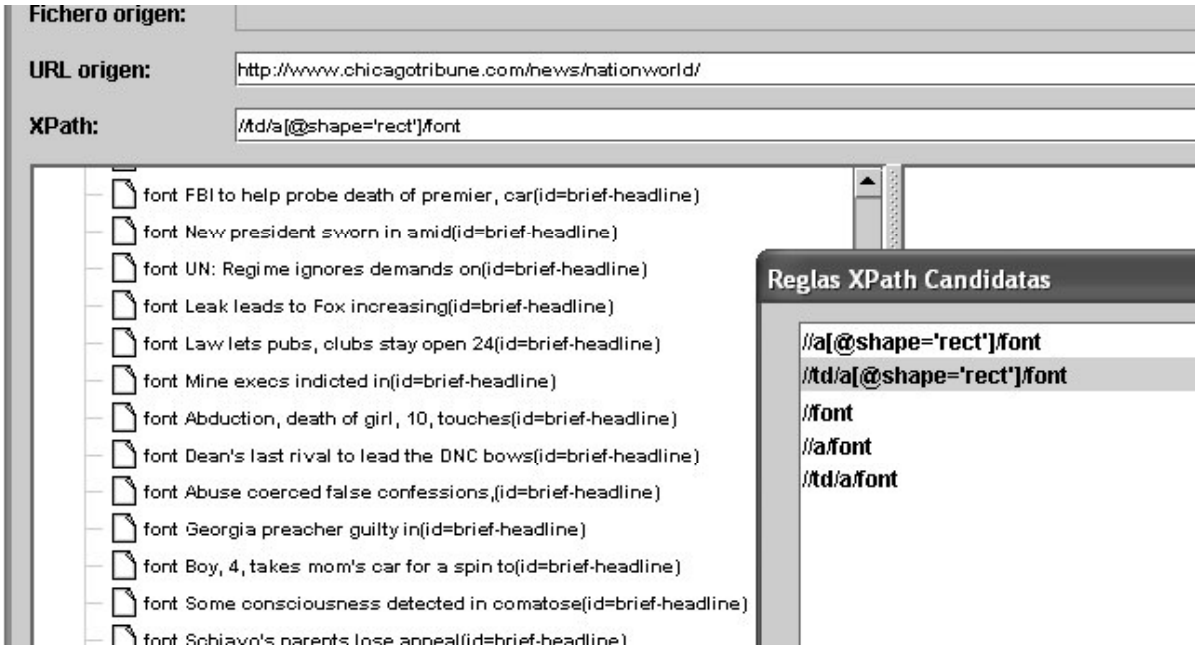


Fig. 5. Resultado de la ejecución de una regla XPath sobre el documento cargado en el visor

Existen alternativas para la verificación de reglas XPath. El problema es que estas alternativas sólo cubren parte de las fases del proceso de creación y prueba.

4 Conclusiones

Si bien es cierto que algunas publicaciones disponibles en Internet proporcionan su información en el formato RSS esto aún no es lo más habitual y son mayoría los contenidos publicados en HTML. Es necesario, por lo tanto, la construcción de extractores con el fin de poder procesar grandes volúmenes de información y la utilización de XPath es una alternativa estándar muy interesante.

En este trabajo se ha creado un sistema de ayuda a la creación de reglas XPath de tal manera que cubra todas las necesidades que han de surgir al desarrollador de extractores, tratando de evitar la utilización de un número amplio de herramientas y cubriendo un hueco, el de la generación automática de reglas XPath, que pueda hacer más accesible esta tecnología.

El sistema desarrollado puede ser utilizado tanto con documentos HTML convertidos a XHTML como directamente con documentos XML. Esto añade la posibilidad de utilizar al sistema para la generación y verificación de reglas XPath para usos diferentes de la creación de extractores de información de la Web.

Agradecimientos

El trabajo en el que se ha basado este artículo ha recibido el apoyo de los proyectos INFOFLEX TIC2003-07208 y SIEMPRE TIC2002-03635 financiados por el Programa Nacional Español de Tecnologías de Información y Comunicaciones.

Referencias

[1] Sitio Web de Cladonia. http://xngr.org/viewer.htm

[2] Amaya. http://www.w3.org/Amaya/

[3] Sitio Web para el Chemical Markup Language. http://www.xml-cml.org/jumbo3/

[4] Sitio Web de X-Smiles. http://www.x-smiles.org/

[5] Sitio Web de Rustemsoft. http://rustemsoft.com/

[6] Sitio Web de XMLFox. http://xmlfox.com/

[7] Sitio Web de XpathVisualizer. http://www.vbxml.com/xpathvisualizer/

[8] Sitio Web de TopXML. http://www.vbxml.com/xsldebugger/

[9] Sitio Web de XML Me. http://www.xmlme.com/XpathTool.aspx

[10] Sitio Web de Purple Technology. http://www.purpletech.com/xpe/index.jsp

[11] Sitio Web de Stylus Studio. http://www.stylusstudio.com/

[12] Sitio Web de Tidy. http://tidy.sourceforge.net/

[13] Sitio Web de JTidy. http://jtidy.sourceforge.net/

[14] Sitio Web de Java. The J2EE™ 1.4 Tutorial. http://java.sun.com/j2ee/1.4/docs/tutorial/doc/JAXPDOM4.html

Generación Automática de Instancias XBRL a partir de Fuentes Propietarias de Información Financiera

Justo N. Hidalgo, Ángel Luengo, Alberto Pan, Ángel Viña

Denodo Technologies
Calle Real, 22 3º
A Coruña 15003, España
Teléfono: +34 981 100 200
E-mail: {jhidalgo, aluengo}@denodo.com

Dpto. de Tecnologías de la Información y la Comunicación
Universidad de A Coruña
Campus de Elviña s/n A Coruña 15071, España
Teléfono: +34 981 167 000
E-mail: {apan, avina}@udc.es

Abstract. *Creation of homogeneous financial reports from several repositories residing at a company's different departments, subsidiaries and/or competitors' publicly available information, is a task which has been eased because of the emergent XML standard for business reporting, XBRL. Nevertheless, this standard is not enough for companies with very heterogeneous repositories, such as relational databases, XML files, spreadsheets, with no single access protocol and, in some cases, semistructured format. This paper proposes the use of the virtual database paradigm for information integration to complement XBRL, and defines a reporting architecture which allows automatic creation of XBRL instances from different databases. By using an ontology-based rule system, the administrator can visually relate those fields from the mediated view to the XBRL Schema tuples, and the application will generate the XBRL instance automatically. The paper also presents a worldwide innovative proof-of-concept which returns instances from a real XBRL taxonomy.*

1 Introducción

La preparación de informes financieros es un proceso crítico para medianas y grandes empresas que ha de ser ejecutado muy cuidadosamente, debido tanto a la importancia que tiene para su credibilidad, como a las dificultades técnicas y burocráticas que entraña. Las razones son tres principalmente. Por una parte, la publicación de un informe financiero en diferentes medios (como por ejemplo la web, archivos legales EDGAR [1] o impresión física) implica en muchos casos la realización de tres subprocesos cuasi-independientes, uno para cada caso. Por otra parte, los datos necesarios para la creación de los informes suelen residir en fuentes de datos muy heterogéneas (bases de datos, hojas de cálculo, ficheros de texto, ..., con lo que el formato de cada fuente es igualmente independiente), remotas y de acceso desigual (p.e. debido a que se encuentran en diferentes departamentos); esto provoca que la creación de los informes se realice manualmente, por lo que la probabilidad de fallos es bastante alta. Por último, aún cuando los informes se generen adecuadamente, no es posible consultarlos para obtener información más detallada o incluso poder aplicar técnicas avanzadas de obtención de conocimiento (p.e. data mining).

Ya existen diferentes estándares de informes financieros, como el Plan General Contable español, el US-GAAP (Generally Accepted Accounting Principles) [2], UK-GAAP [3], IFRS (Internationally Financial Reporting Standards) [4], pero no de forma, sino de fondo, pues no establecen en ningún caso una estructura fija de relación entre elementos, ni definen reglas de cálculo y/o presentación.

La aparición del estándar XBRL (XML Business Reporting Language) [5] ha supuesto un avance muy importante en la automatización y estandarización de la creación de informes financieros, convirtiéndose en la base a partir de la cuál parte el resto de formatos de presentación. Además, XBRL es un lenguaje consultable, por lo que potencia la capacidad de obtención de información de valor añadido de manera automática.

Sin embargo, XBRL por sí sólo no soluciona todos los problemas descritos. La generación de instancias XBRL a partir de las fuentes de datos es un desafío que el estándar no puede resolver. Aunque existen soluciones propuestas por diferentes centros y empresas [6], [7], [8], ninguna ha sido capaz de automatizar plenamente el proceso de captura.

En este artículo se presenta y describe una solución holística de definición, integración, extracción de datos y generación de instancias XBRL conformes al estándar. Este sistema, que se basa en tecnologías y técnicas de integración de información, bases de datos virtuales y extracción automática de información, es la primera prueba de concepto de este tipo de soluciones en el mercado XBRL.

El resto del artículo se divide de la siguiente manera. La sección 2 introduce XBRL, sus puntos fuertes y qué deficiencias tiene con respecto a la integración de información. La tercera sección describe las soluciones existentes en cuanto a la obtención de datos que sirvan de fuente a la generación de instancias XBRL, de manera que la sección 4 propone un enfoque basado en el paradigma de Mediación (o Base de Datos Virtual). Por último, el

apartado de conclusiones resume los objetivos alcanzados con la implementación de esta arquitectura, conclusiones y el trabajo futuro a realizar.

2 El Estándar XBRL

XBRL (eXtensible Business Reporting Language) [5] es un lenguaje para la comunicación electrónica de datos financieros y de negocio. Es un estándar abierto creado por XBRL International, y que pertenece a la familia de los lenguajes XML [9], el cual se ha convertido en el estándar para la comunicación de información entre negocios y en Internet.

La idea de XBRL es muy sencilla. En lugar de tratar la información financiera como si de un texto plano se tratara (como puede ser una página estándar de Internet o una página impresa), lo que se hace es proveer etiquetas para identificar cada elemento del documento. Estas etiquetas se generan siguiendo el estándar XML [9]. XBRL se define para cumplir con las necesidades de negocios e información financiera, lo cuál hace que estos documentos sean fácilmente interpretados con un ordenador. Es decir, permite el procesamiento automático de información financiera mediante software evitando tediosos procesos manuales. Gracias a las etiquetas, una aplicación puede reconocer la información contenida en un documento XBRL, pudiendo así seleccionarla, analizarla, guardarla, intercambiarla con otras aplicaciones/ordenadores y desplegarla de manera automática en multitud de maneras distintas según las necesidades del usuario. Estas etiquetas proporcionan, además, información sobre el objeto que representan (p.e. si el objeto representa un porcentaje, una moneda, etc.). En resumidas cuentas, XBRL minimiza los tiempos en el tratamiento de información financiera, reduciendo errores. Por otra parte, XBRL muestra además como los objetos están relacionados unos con otros utilizando XLink [10], Namespaces [11] y XPath [12], permitiendo así saber cómo son calculados. Esto es muy importante ya que la generación automática de informes se basa en esta característica.

Otra de las grandes ventajas de XBRL es que es fácilmente extensible pudiendo así compañías y organizaciones adaptarlo para cumplir requerimientos específicos.

XBRL se ocupa sólo del *reporting*, por lo que no entra en contienda con otros estándares tales como: XML/EDI [13], OFX [14], IFX [15], Fix [16], FinXML [17], etc., que son protocolos orientados a transacción.

XBRL no sólo utiliza las tecnologías XML para estructurar información. XBRL permite crear ontologías, ya que gracias a XMLSchema puede determinar una taxonomía concreta, mientras que la utilización de XLink, XBRL origina una serie de ficheros relacionados con el "core", denominados XBRL Linkbases; cada uno de estos ficheros tiene una tarea concreta y desacoplada del *core*. Es decir, que el *core* XBRL sólo determina la relación estructural entre sus elementos, mientras que los Linkbases definen otras relaciones, como por ejemplo de presentación, de cálculo (lo cuál permite confirmar la validez de elementos agregados) o de etiquetado, vital para las capacidades de internacionalización de la estructura: el "core" no define la representación de sus valores en diferentes idiomas, sino que se delega a diferentes Label Linkbases, tantos como idiomas se requieran. Resumiendo, estos Linkbases determinan las reglas que toda ontología tiene sobre su taxonomía.

La aceptación de XBRL como estándar de generación de informes financieros soluciona gran parte de los problemas existentes en la actualidad y de una manera no demasiado intrusiva. XBRL no redefine las normativas actuales, tanto internacionales (p.e. la ya comentada IFRS [4]) como locales (p.e. US-GAAP [2], UK-GAAP [3] o el Plan General Contable español), sino que las reescribe al formato XML establecido. Sin embargo, no es suficiente. La generación previa de estos informes XBRL es ya de por sí un proyecto muy ambicioso para empresas cuyos datos se encuentran distribuidos por diferentes repositorios de información de multitud de departamentos. La utilización de mecanismos manuales o de automatizaciones ad-hoc conlleva una alta probabilidad de errores.

3 Enfoques de integración

Además de la integración manual, la generación de informes financieros puede realizarse desde otros enfoques.

El primer enfoque implica la creación de un sistema materializado o Base de Datos Universal, es decir, la creación de un único repositorio unificado a partir del cuál sea muy sencillo construir el informe financiero en XBRL u otro formato. Consiste en la creación de una nueva base de datos cuyo esquema unifique el de todas las bases de datos a integrar. Las antiguas bases de datos dejan de funcionar y todas las operaciones pasan a realizarse contra la nueva, incluidas las de escritura. Aunque obviamente es un enfoque muy eficiente, tiene el grave inconveniente de que todas las aplicaciones que estaban funcionando sobre las viejas bases de datos, dejan de funcionar y, por lo tanto, es necesario reprogramarlas o desarrollar otras nuevas. Como consecuencia, posiblemente también sea necesario realizar programas de formación en el uso de las nuevas aplicaciones, para los usuarios de las viejas. Además, será necesario hardware nuevo más potente que el anterior para que el sistema sea capaz de realizar todas las operaciones que antes realizaba por separado cada una de las bases de datos. Por lo tanto, las migraciones son largas y muy caras.

Otro inconveniente importante es que si en el futuro es necesario integrar otra base de datos, todo el sistema puede verse afectado, lo que compromete gravemente la extensibilidad del enfoque. Como ventaja, este enfoque será más eficiente que el resto, ya que, al fin y al cabo, una vez realizado, no deja de ser una base de datos convencional.

El segundo enfoque es el utilizado por los sistemas de Data Warehouse. Al igual que en el enfoque anterior, se construye una nueva base de datos unificada que contiene datos de todas las bases a integrar. La diferencia fundamental es que, en este caso, las bases de datos originales no desaparecen sino que siguen funcionando. Los datos del almacén se utilizan sólo para propósitos de lectura (normalmente para realizar algún proceso de análisis o minería de datos sobre los mismos), mientras que el resto de procesos siguen funcionando de la misma manera que antes de la creación del almacén. Tiene la desventaja de que es necesario actualizar el almacén cuando los datos cambian. Como esto no siempre es posible o fácil de realizar, a menudo los datos se encuentran sin actualizar.

Además, no siempre será posible copiar todos los datos de los sistemas originales en el almacén. Por ejemplo, en el caso de una interfaz de consultas web –caso muy típico en la actualidad-, ésta no siempre permite extraer todos los datos. Un Web Service de información pública de competidores permitirá buscar en su catálogo por el C.I.F. de la empresa, pero no permitirá acceder a una lista de todas las empresas de su catálogo, de manera que éste pueda ser almacenado completo en el sistema de integración. Otra razón por la que puede no ser posible copiar todos los datos es que, sencillamente, su volumen puede ser demasiado grande para ser transferido por la red en un tiempo razonable.

Los sistemas tradicionales obtienen los datos necesarios para la generación de informes de un conjunto de fuentes, generalmente relacionales o data warehouses. En la mayor parte de las ocasiones, este tipo de soluciones presentan inconvenientes muy serios para los modernos entornos de negocio:

- Debido a su funcionamiento off-line, no constituyen un esquema válido cuándo se requiere que los datos integrados estén actualizados en tiempo real, que en muchas ocasiones es el caso de los informes financieros. El acceso a la información corporativa en tiempo real permite disminuir enormemente los tiempos de reacción en la toma de decisiones, y ha sido identificado por diversos estudios de consultoras como Aberdeen y Gartner Group como un factor clave en la competitividad de las empresas.

- En un entorno competitivo como en el que se encuentran la mayor parte de las empresas en la actualidad, los indicadores utilizados en los informes financieros han de utilizar tanto fuentes internas de diferentes departamentos –repositorios heterogéneos controlados de manera dispersa y en áreas de trabajo con mucha presión, lo cuál impide una integración intrusiva adecuada- como elementos públicos externos –páginas web de competidores, información pública sobre las áreas de actividad de interés, etc.-. Este tipo de fuentes no admite intrusividad –i.e. un agente ejecutándose en la máquina de la cuál proceden los datos que preprocese esa información-, por lo que los sistemas tradicionales son inútiles.

- Requieren la construcción de un nuevo gran servidor central que contendrá un gran volumen de datos, lo cuál es caro, lleva a proyectos muy largos y es poco escalable

El último enfoque, presentado en este artículo se basa en la utilización de mediadores o bases de datos virtuales. El sistema mediador integra las fuentes de datos estructuradas y semiestructuradas, sacando provecho de la estructura parcial de estas últimas para permitir, también sobre ellas, la realización de consultas precisas escritas en lenguajes de consulta estructurados como los utilizados sobre bases de datos convencionales (y, por lo tanto, estructuradas). Este sistema no utiliza el paradigma materializado, sino el virtual. La información es obtenida en tiempo real de las fuentes remotas, de una manera no intrusiva –es decir, sin necesidad de agentes en las máquinas que lo soportan-, e integradas en una vista global generada localmente. Las ventajas con respecto a las iniciativas descritas anteriormente son evidentes:

- No es necesario realizar ningún proceso de reingeniería de las aplicaciones utilizadas sobre las bases de datos originales.
- El nuevo hardware necesario es mínimo, ya que el sistema de bases de datos federadas sigue delegando el grueso del trabajo en las bases originales.
- Los datos están permanentemente actualizados y están todos disponibles, con lo que pueden utilizarse para cualquier tipo de aplicación y no sólo aquellas relacionadas con la minería de datos sobre históricos. Si el administrador lo desea, en algunos casos le será permitido mantener caches de los datos de las bases originales, pero eso será siempre una decisión que podrá tomar en función de la naturaleza de la aplicación a construir.

La generación de informes financieros requiere en la mayoría de los casos, el acceso a fuentes estructuradas y semiestructuradas para la obtención de datos que suelen cambiar con frecuencia. Es por ello por lo que el siguiente apartado describe una solución completa de Generación de Instancias XBRL basada en un sistema mediador de datos.

4 Arquitectura de la Solución

El objetivo de esta arquitectura es demostrar que XBRL es una solución útil y fiable en el mercado de información financiera, y que la utilización de un mediador permite obtener esos informes a partir de fuentes preexistentes con formatos y protocolos de acceso heterogéneos.

La Figura 1 muestra esta arquitectura. Se observa cómo la solución de integración se divide en dos partes fundamentales. Por un lado el mediador, cuya estructura se basa en [18] y sigue el paradigma Global-As-View [19], abstrae la información integrada por medio de tablas virtuales –vistas de los datos de las fuentes- siendo por tanto el responsable de responder a las peticiones realizadas por el usuario, mediante consultas VQL (Virtual Query Language) [20], que es un lenguaje muy parecido a SQL. Por el otro, el mediador delega el acceso a las fuentes remotas a los "*wrappers*" o envoltorios, que conocen tanto el protocolo de acceso a cada fuente, como el modo de extracción de datos semiestructurados (p.e. de fuentes web) y estructurados (p.e. bases de datos relacionales); una vez obtenidos, estos *wrappers* devuelven los datos convenientemente estructurados según la definición de la tabla virtual, al mediador.

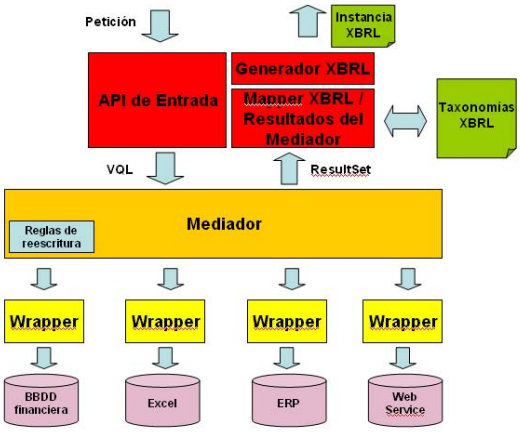


Figura 1: Arquitectura Genérica de Mediación

Una vez el mediador responde la consulta concreta del usuario, ésta no es devuelta directamente, sino que es utilizada por el módulo de generación XBRL para crear la instancia XBRL requerida. Para ello, se realizan dos acciones básicas:

- Obtención de la/s taxonomía/s XBRL necesarias a partir de las cuáles realizar el *mapping* con el resultado devuelto por el mediador. Las reglas de *mapping* se han generado previamente de manera declarativa, y los algoritmos creados para ello están fuera del alcance de este artículo, pero merece la pena comentar la importancia de que estas relaciones se definan declarativamente, de manera que la información de relación entre las vistas VQL creadas y los datos almacenados en las taxonomías de interés para el usuario pueda ser generada visualmente, y una sola vez.

- Una vez ese *mapping* se ha realizado, se crea una instancia XBRL conforme a esa/s taxonomía/s, y poblada con los datos obtenidos por el mediador. Esta instancia se crea siempre en tiempo real, pues al contar con un mediador que permite obtener los datos origen rápidamente, el resultado de la aplicación es inminente. Aún así, el mediador puede hacer uso de su cachè para que los datos que no cambian de una manera frecuente no sean accedidos constantemente, por lo que se mejora la eficiencia en los tiempos de respuesta a consultas.

Las ventajas de este sistema son evidentes, y se describen a continuación. Por una parte, la creación de las vistas virtuales, aunque ha de realizarse manualmente, sólo ha de hacerse una vez, y mediante herramientas gráficas que simplifican su generación. En el caso de fuentes relacionales, web services, ficheros XML o ficheros Excel, los *wrappers* que acceden a su información se crean automáticamente en cuanto sus datos de acceso (p.e. protocolo, driver, usuario y contraseña en el caso de bases de datos relacionales accesibles mediante JDBC u ODBC) se introducen en el sistema, mientras que si la fuente es web, es necesario un trabajo adicional de generación visual de la secuencia de navegación (muy necesario en el caso de webs transaccionales, en las que se requiere una navegación previa a través de formularios de autentificación, de búsqueda, menús y/o enlaces) y de generación del programa de extracción de información, para convertir el conjunto de datos semiestructurados en información estructurada. Estos pasos se realizan mediante otra interfaz gráfica. Una vez las fuentes de datos han sido adecuadamente modeladas, el administrador puede crear las vistas compuestas que desea de manera visual. Estas vistas son almacenadas en el mediador, de manera que pueden utilizarse las veces que sean necesarias.

La segunda ventaja es que el sistema no requiere que el administrador tenga conocimiento acerca del estándar XBRL que, aunque muy útil, no está pensado para que sea tratado y procesado manualmente. La generación de reglas se realiza entre abstracciones de XBRL que sólo muestran la taxonomía almacenada, no los detalles de especificación de XBRL, XML, XLink, etc.

Como tercera ventaja, y relacionada con la anterior, debido al desacoplamiento entre la estructura del mediador y las taxonomías, relacionadas a través de un conjunto de reglas definibles visualmente, es muy sencillo, casi trivial, poder generar distintas instancias XBRL dependientes de diferentes taxonomías que utilizan conjuntos no disjuntos de las mismas fuentes.

Además, debido a la modularización de las reglas de *mapping*, si una segunda taxonomía utiliza las mismas reglas sobre una fuente en particular –y no obligatoriamente sobre el resto-, este subconjunto de reglas puede ser reutilizado. Empíricamente se comprueba que suele ser un caso habitual.

5 Prueba de Concepto Construida

Para demostrar la capacidad de la arquitectura descrita en el apartado anterior, se ha construido una prueba de concepto. Se definen dos fuentes heterogéneas de utilidad real y se realiza una integración de esas fuentes mediante el mediador, y la generación de instancias XBRL de sus resultados, utilizando la taxonomía FLIPA [21].

La taxonomía FLIPA pertenece a los tipos de taxonomía XBRL-CRAS (Credit Risk Assessment Services). Este tipo de taxonomías permiten generar informes que faciliten la gestión de riesgos.

Supóngase que una entidad X necesita ser asegurada. La entidad Y, encargada de hacer el seguro, utilizará los informes generados a partir de la taxonomía FLIPA para obtener información útil sobre la entidad X y calcular así los riesgos que asume al asegurarla.

Utilizando documentos XBRL generados en base a la taxonomía FLIPA sobre ella, el proceso se puede automatizar haciéndolo más sencillo, rápido y fiable.

La taxonomía FLIPA permite, por tanto, *reporting* tanto de datos financieros como de datos que no lo son, no solo de una entidad/empresa sino también de sus filiales, lo cual permite un análisis más exhaustivo sobre la situación de la empresa.

Los documentos XBRL generados a partir de esta taxonomía están principalmente destinados a sistemas de información que sean capaces de automatizar la gestión de riesgos y la toma de decisiones a partir de ellos.

La información que contiene el documento XBRL no tiene por que provenir de la entidad a la que afecta, sino que puede provenir de distintas fuentes y proveedores, lo cuál fue una de las razones primordiales por las que se eligió esta taxonomía como base a partir de la cuál crear esta prueba de concepto. Además, la taxonomía FLIPA cuenta con la particularidad de que la definición de conceptos, el *core* FLIPA, se encuentra dividido en dos ficheros distintos, el FLIPA Core y el FLIPA Datatype, siendo la segunda una extensión de la primera. El FLIPA Datatype se encarga de definir los valores reales que pueden tomar los conceptos abstractos definidos en el FLIPA Core. Esta grado de separación permite que al hacer algunas modificaciones –p.e. internacionalización- no sea necesario modificar el fichero principal, el FLIPA Core, sino tan solo el Datatype. Esta peculiaridad permite probar de una manera más exhaustiva la genericidad de la solución propuesta.

Las fuentes que se han utilizado son las siguientes:

- Un Web Service que almacena información financiera sobre empresas, accesibles a través de su CIF.

- Una base de datos relacional accesible mediante ODBC, que contiene datos financieros y estadísticos sobre diferentes sectores de actividad. Una empresa puede relacionarse con uno o más sectores.

Ambas fuentes contienen información financiera real de empresas españolas. Estos dos tipos se han seleccionado pues responden adecuadamente a los factores de prueba de este sistema:

- Repositorios con datos reales, y con la misma estructura que en situaciones reales de la industria.

- Acceso a bases de datos relacionales, uno de los tipos de fuente más utilizados en el entorno financiero.

- Acceso a WebServices, tanto por ser uno de los tipos de fuente que más crecimiento está teniendo, como por su débil estructura y gran jerarquización, lo cuál permite evaluar los algoritmos y técnicas de relación entre las vistas virtuales del mediador y la instancia XBRL objetivo.

La Figura 2 muestra la relación entre los diferentes componentes de la prueba de concepto.

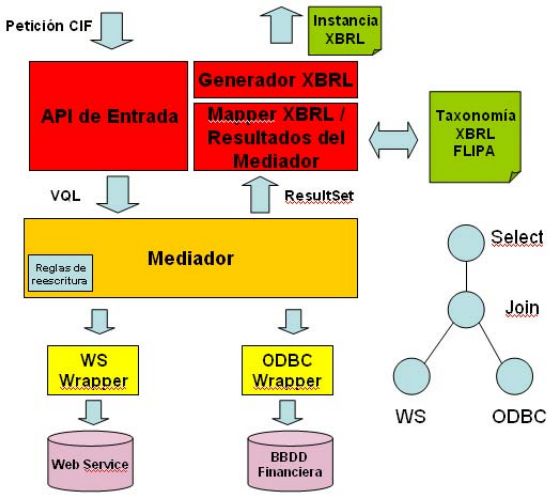


Figura 2: Arquitectura de la Prueba de Concepto

El flujo de información es el siguiente:

1. El usuario dispone de una página web desde la cuál se puede realizar una consulta a través del CIF sobre cualquier empresa. Este es el parámetro de entrada.

2. El Generador XBRL, con el CIF, invoca al mediador realizando una consulta mediante VQL (Virtual Query Language); esta consulta podrá ser:

   - Una consulta única a la información de la empresa almacenada en el WebService.

   - Un *join* a la fuente anteriormente mencionada, y a la de información sectorial sobre áreas de actividad empresarial. Esta opción muestra con mayor claridad la capacidad de la arquitectura de integración y generación XBRL.

3. El WebService devuelve un XML con un formato determinado que el *wrapper* convertirá a tabla virtual mediante las reglas de reescritura del mediador.

4. Por otra parte, la otra fuente, ODBC, devuelve sus datos como un conjunto de filas que es convertida a tuplas de una tabla virtual del mediador.

5. El componente de generación XBRL realiza el *mapping* final entre la información publicada por el mediador y la taxonomía XBRL FLIPA [21].

6. El componente devuelve el documento instancia XBRL al usuario. La Figura 3 muestra un extracto de una instancia devuelta.

En la descripción de esta prueba de concepto se ha dado por hecho que las reglas de relación han sido creadas previamente.

Figura 3: Extracto de Instancia XBRL resultado

6 Conclusiones

En este artículo se ha presentado una arquitectura innovadora que hace uso de mediadores de información heterogénea como base de acceso estructurado a datos financieros dispersados en diferentes repositorios, para la creación automática y en tiempo real de instancias XBRL que cumple un conjunto de taxonomías determinado. Esta arquitectura, además, ha sido validada por una prueba de concepto sobre una taxonomía real como es FLIPA, que ha demostrado por primera vez la posibilidad de que los sistemas mediadores formen parte de soluciones de este tipo, de carácter industrial y comercial. Las dos fuentes utilizadas demuestran la versatilidad del sistema en cuanto a heterogeneidad de formato (estructurado en el caso del repositorio relacional, jerarquizado y semiestructurado en el caso del WebService). Por otra parte, el sistema mediador ya se ha utilizado previamente en gran cantidad de implantaciones, con lo que su escalabilidad y tolerancia a fallos ha sido demostrada. La solución presentada en este artículo se ha convertido en el paso previo al desarrollo de un producto vertical comercial.

7 Trabajo Futuro

El trabajo a realizar tras la elaboración de este prototipo es, desde el punto de vista industrial, la mejora de usabilidad en las herramientas de generación visual de relaciones entre las vistas virtuales y las diferentes taxonomías XBRL de interés.

En cuanto al aspecto de investigación, creemos que este proyecto abre nuevas vías en la generación de información integrada a partir de fuentes heterogéneas, de manera que los datos devueltos cumplan con una o varias ontologías determinadas en tiempo real y de manera semiautomática. Un siguiente punto de investigación por tanto será la generalización de la prueba de concepto para la obtención de un producto independiente de características internas de XBRL y abierto a cualquier taxonomía creada mediante XML.

Referencias

[1] EDGAR. The Electronic Data Gathering, Analysis and Retrieval System. http://www.sec.gov/edgar.shtml

[2] US-GAAP. United Sates - Generally Accepted Accounting Principles. Taxonomy Framework-2005-02-28.

http://www.xbrl.org/taxonomy/us/fr/gaap/ci/2002-10-15/

[3] UK-GAAP. United Kingdom - Generally Accepted Accounting Principles. http://www.xbrl.org/uk/fr/gaap/ci/2004-05-15/uk-gaap-ci-2004-05-15.doc

[4] IFRS. International Financial Reporting Standards (IFRS), General Purpose Financial Reporting for Profit-Oriented Entities (GP), 2004-06-15. http://xbrl.iasb.org/int/fr/ifrs/gp/2004-06-15/

[5] XBRL 2.1. XML for Business Reporting Language Specification, 2003-12-31. http://www.xbrl.org/SpecRecommendations/

[6] Microsoft Office Solution Accelerators. http://www.microsoft.com/office/solutions/xbrl/default.mspx

[7] Fujitsu XBRL. http://www.fujitsu.com/global/services/software/interstage/products/xbrlprocessor/xbrl.html

[8] DecisionSoft, XBRL Solutions. http://www.decisionsoft.com/xbrl/

[9] Extensible Markup Language (XML) 1.0 (Third Edition), W3C Recommendation 04 February 2004. http://www.w3.org/TR/2004/REC-xml-20040204/

[10] XML Linking Language (XLink) Version 1.0. W3C Recommendation 27/6/2001. http://www.w3.org/TR/xlink/

[11] Namespaces in XML World Wide Web Consortium,14-January-1999. http://www.w3.org/TR/REC-xml-names/

[12] XML Path Language (XPath) Version 1.0, W3C Recommendation 16 November 1999. http://www.w3.org/TR/xpath

[13] Extensible Markup Language (XML) / Electronic Data Interchange (EDI). http://www.geocities.com/WallStreet/Floor/5815/

[14] OFX. Open Financial Exchange. http://www.ofx.net/ofx/ab_main.asp

[15] IFX. Interactive Financial eXchange. http://www.ifxforum.org/ifxforum.org/standards/index.cfm

[16] FIX. Financial Information eXchange. http://www.fixprotocol.org/what-is-fix.shtml

[17] FinXML. Financial XML. http://www.finxml.org/Information/Information.asp

[18] Wiederhold, G. Mediators in the Architecture of Future Information Svstems. IEEE Computer, March 1992, pp. 38-49

[19] Cali, A., Calvanese, D., De Giacomo, G., and Lenzerini, M. On the expressive power of data integration systems. Proc. of the 21st Int. Conf. on Conceptual Modeling (ER), 2002, pp. 338-350.

[20] Pan, A., Raposo, J., Álvarez, M., Montoto, P., Orjales, V., Hidalgo, J., Ardao, L., Molano, A., Viña, A. The DENODO Data Integration Platform. Proc. of the 28th International Conference on Very Large Data Bases (VLDB 2002)., pp. 986-989.

[21] XBRL CRAS (XBRL Credit Risk Accessment Services) FLIPA. http://demo.ubmatrix.com/cras/XBRL%20CRAS_files/ubm_AboutCRAS.htm

A Dynamic Web Programming Methodology to measure the impact of subjective factors on the perceived VoIP speech quality

Rafael Estepa, Antonio Estepa, Juan Vozmediano
Area de Ingeniería Telemática. Universidad de Sevilla
Escuela Superior de Ingenieros.
C/Camino de los descubrimientos s/n. 41092 Sevilla
E-mail: {rafa,aestepa,jvt}@trajano.us.es

Abstract *Automation of opinion tests to evaluate the QoS in VoIP can enhance the tayloring of well-established predictive models such as the E-Model. Two small surveys to adjust this model have been performed with the help of a dynamic Web server. This test methodology can be easily scaled up. The results show that VoIP can be highly valuable in trade off new wideband codecs, new environment and low cost especially in specific groups of population like young people.*

1 Introduction

The quality of service (QoS) in telephony is a subjective concept involving the user's perception of the underlying transmission system. This inherent subjectivity leads to perform opinion tests on speech and conversational scenarios to properly evaluate the QoS level.

The high cost and time required to perform these tests impelled the development of methods to predict the average user opinion from objective, measurable transmission parameters.

Among these methods, the E-Model [1] seems to be flexible enough to be taylored to assess the QoS of VoIP services. The E-Model outcome is a quality score, named Transmission Rate factor (R), ranging from 0 to 100, derived from an additive expression that takes into account several physical, measurable parameters plus a correction term A. The latter, known as expectation factor, models the decrease in R that the user tolerates in exchange for other non-measurable, desirable features.

$$\begin{aligned} R &= R_0 - I_s - I_d - I_e + A \\ &= M + A \end{aligned} \quad (1)$$

The R factor can be directly mapped to the Mean Opinion Score (MOS), traditionally used to rate telephony quality in a scale from 5 (excellent) to 1 (bad). This relationship is given by:

$$\begin{aligned} MOS &= 1 + 3.5\ 10^{-2} R \\ &\quad + 7\ 10^{-10} R(R-60)(100-R) \end{aligned} \quad (2)$$

The tuning of the E-Model to the VoIP scenario is in progress [2], but the translation of subjective parameters into the psychological scale have to be validated through opinion tests. The proper methodology is described at [3] and outlined at [4].

For those conditions in which E-Model had been validated, deviations between the predicted R value given by the E-Model and the average user rating can be utilized to adjust the A factor. The adjustment of E-Model factors by using objetive method (like PESQ) does not allow the delay inclusion in test condition. Nevertheless, the delay is one of the most important impairment in VoIP, which forces us to use opinion test.

We have performed two small surveys to tune this factor, with the aim of designing and testing a reusable and scalable methodology to perform wider surveys. The proposed system allow the centralized processing of the opinion test when compared with other similar systems (like Xcorpus, from Loria). Both surveys are also intended to contribute to include the price influence in the perceived conversation quality.

The design and methodology are described in sections 2 and 3 respectively. The results obtained can be examined in section 3. Section 4 analyzes the results and section 5 concludes the paper.

2 Test Design

2.1 Listening Test

For the listening test, four talkers (two males and two females) provided a total of twelve recorded speech samples according to [3]: each sample consisted of two sentences 8 to 10 seconds long, with a separation pause of 0.5 seconds. Two similar pauses were inserted at the beginnig and end of

each sample.

These speech samples were coded with a narrowband G.711 and a wideband G.722 codec. Eight reference conditions were established as a baseline for comparison to different test rooms. These reference conditions were recorded adding controlled and known degradations to the source material with a Modulated Noise Reference Unit, and sampled at 16KHz.

The *ITU-T Software Tool Library* provided packages for equalization, sampling and coding of the recorded material.

After a first test run without price considerations, two more runs were made, telling the user that the rate was either half the price or same price than POTS dominant provider. Since we have used two different codecs, this yields a number of 6 voting conditions (without considering the 8 reference conditions).

Since we recorded twelve speech samples and we present fourteen different conditions for each sample. Thus we have a total of 168 sample files to be played.

2.2 Conversational Test

The second test consisted of a series of conversations using a PC-based telephone set. Following [3] and [4], the conversations were about a card game in which each users had an arbitrary set of cards and a common ordination had to be agreed. Each conversation lasted 180 s. The minimund delay that any conversation will experience due to the network and terminals was experimentally measured, obtaining a result of 200 ms. An extra incremental delay of 0, 400 and 600 ms was inserted in between of each conversation.

Price rates were also indicated before every test. The codec was G.723, as preferred by the ITU-T [5], configured to 6.3 Kb/s and 3 frames per packet.

3 Methodology and Test Results

The testbed was set up in three separate rooms (Fig. 1). The listening test took place in room A, equipped with four listening booths. According to [3], two different persons were holding a conversation between rooms B and C for the conversational tests.

The network under test was a dedicated Fast Ethernet, which guaranteed negligible packet loss and delay. Each test was conduced by a dedicated Web server running Apache and a Postgres database. The dynamic web pages were programmed with PHP as language. The workflow of the web pages presented guided the user through the listening process and collecting the scores. The web server used for the conversational test was also able to automatically insert the required delays thanks to the Nistnet [6] network emulator.

One of the web servers was used for recruiting participants from a group of engineering students between 19 and 21 years old. We are aware that this introduces a technology-benevolent bias in the results.

The first step in each test included identifying the participant, giving general intructions, and adjusting the volume optimal level using test samples. The participants were informed about the VoIP technology that was being used. This new environment (PC terminals and network) can represent a tradeoff advantage for this population segment.

The test condition were randomly chosen, always meeting the constraint that tests without cost references were played first. After each sentence or conversation, the participant had 5 seconds to rate the quality in a scale of five values. Averaging over all registered scores in the experiment resulted the MOS.

Timing for every vote was also recorded, allowing supervision and scanning for invalid votations (i.e. response times too short or repetitive constant scores). A new test could re-use the existing system by just storing the new sample files in a specific directory and some minor modifications of the web pages.

For the listening test, the set of 168 coded speech samples were rated by 55 participants, giving 660 scores for each one of the fourteen listening conditions. The participants only had to click in a browser for listening and voting. The Web pages were dynamically generated from the available sample files. The results are summarized in Fig. 2, where the average rating, and the 95% confidence intervals are depicted for each condition.

Users penalize cost with a decrement of up to 20% of the perceived quality, but the G.711 codec is rated 16% lower than G.722. This confirms that users can tolerate higher prices in trade off higher quality. The effect of the wideband codecs should be taken into account in the I_e factor of the model. Some other studies present test results for the G.722 codec [7, 8], but further work is needed to assess the I_e values to be included in the E-Model.

For the conversational test, a total of 38 valid participants were chosen for this test. Fig. 3 shows the average and the 95% confidence interval for each condition. A star-mark shows also the MOS predicted by the E-Model. Again, conversations without pricing references were held first.

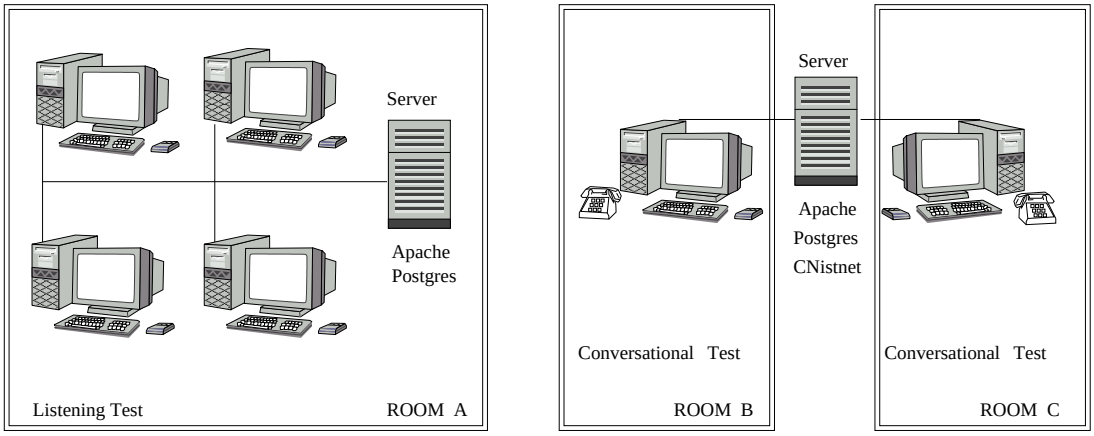


Figure 1: Testbed and servers used for conversational test (rooms B and C) and listening-only (room A)test.

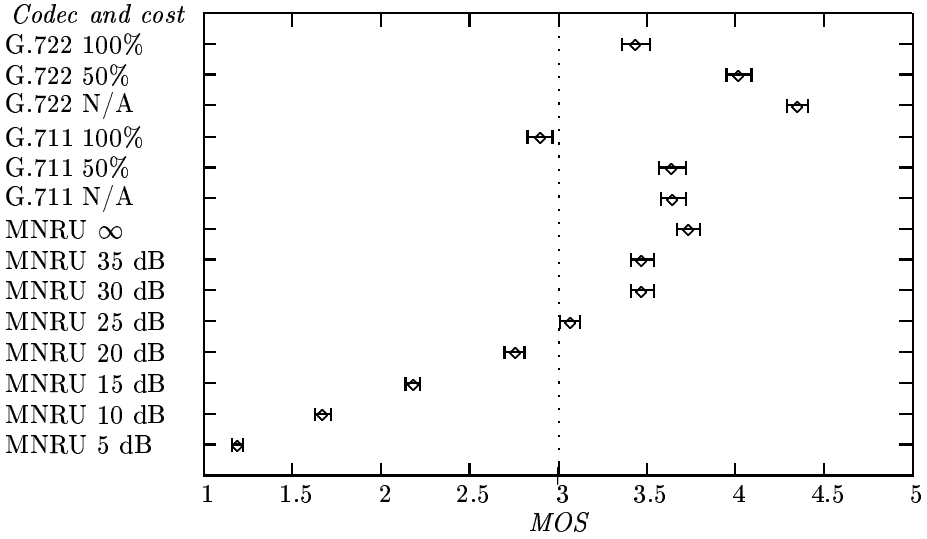


Figure 2: Listening Test results (MOS)

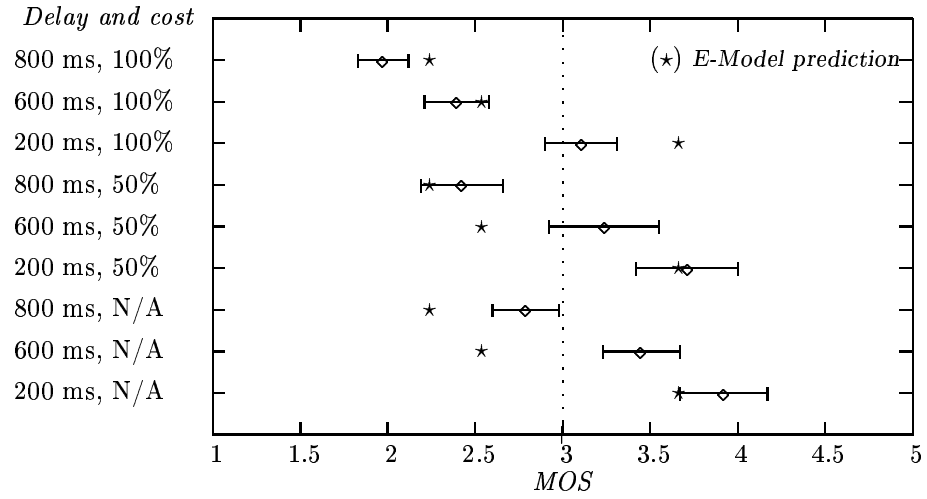


Figure 3: Conversational Test results (MOS)

4 Tunning of the A factor

To tune the A factor it is necessary to compare the results above and the predicted MOS value obtained from the output of the E-Model. It's well known that E-Model accuracy depends on the test conditions. For example, the packet loss impairment is not addresses in the E-Model results for bursty losses (as usual in the Internet). Those conditions in which E-Model has a lack of accuracy are avoided in the test.

This requires measuring from the testbed input parameters for the E-Model such as delay, echo and sidetone. Only the conversational test was considered, because it includes all the objective and subjective factors in a VoIP conversation.

We believe that the A factor can be split into two independent subfactors whose additive effect includes the use facility of VoIP and its cost advantage. The former can be justified because some users may positively value the use of a PC or laptop as a telephone set connected to the Internet. The latter can be related to the widespread perception that VoIP is invariably cheaper than POTS, as a consecuence of the lower provision cost from the provider's perspective.

$$A = A_{access} + A_{cost} \tag{3}$$

>From Fig. 3, averaging the differences from conditions with same cost references and the corresponding E-Model predicted values, and converting them to the R scale, we get:

$$A(N/A) = 16.313 \tag{4}$$
$$A(50\%) = 8.790 \tag{5}$$
$$A(100\%) = -9.303 \tag{6}$$

As A_{access} is included and known, substracting from above:

$$A_{access} = 16.313 \tag{7}$$
$$A_{cost\,=\,N/A} = 0 \tag{8}$$
$$A_{cost\,=\,50\%} = -7.52 \tag{9}$$
$$A_{cost\,=\,100\%} = -25.62 \tag{10}$$

The full A factor proposed is the polymonial fitting:

$$A = A_{access} + A_{cost} \tag{11}$$
$$= 16.313 - 11.51r - 14.1r^3 \tag{12}$$

where r is the cost ratio to POTS.

5 Conclusions

This paper has proposed a flexible and scalable methodology to perform VoIP opinion tests. The whole survey was conduced by a Web server where Web pages were dynamically generated from the available sample files. The results were gathered and processed automatically. Thanks to the use of databases and dynamic web programming, dynamic instructions can be used for each condition. This can be very useful in case of testing any pure subjective factor (like cost). This methodology also can perform some control tasks that traditionally had to be done manually. We found that this methodology can be easily scaled up with the number of participants or test conditions, by just adding new condition files.

Also some tuning of the E-Model have been proposed to fit it to VoIP. The cost ratio to POTS influences the final quality perceived by the users and it should be taken into account in the model, because costs still represents a big difference between POTS and VoIP Telephony and it is likely that convergence in prices doesn't come soon.

The new environment with different terminals and the knowledge of the underlying transmission technology makes a positive influence in the participants, taken from an homogeneous group of students.

The use of wideband codecs influences positively the users opinion and maybe can be the key issue for the quick growing of VoIP technology in the future.

Agradecimientos

The work leading to this article has been partly supported by CICYT and the EU under contract number 1FD97-1003-C03-03.

References

[1] *Speech Communication Quality from mouth to ear for 3,1 KHz handset telephony across networks*, ETSI Std. ETR250.

[2] *Provisional planning values for the equipment impairment factor Ie*, ITU-T Std. G.113 Appendix I.

[3] *Methods for subjective determination of transmission quality*, ITU-T Std. P.800, Aug. 1996.

[4] *Telephonometry Manual*, ITU-T.

[5] *Packet-based multimedia communications system*, ITU-T Std. H.323v4, Nov. 2000.

[6] [Online]. Available: http://www.itl.nist.gov/div892/itg/carson/nistnet/

[7] A. Duric, "Speech coders - a VoIP perspective," in *Workshop on QoS in Next Generation Networks*. ETSI, Mar. 2002.

[8] *Evaluation of subjective quality in wideband codecs*, ITU-T Std. H.323v4, Feb. 1996.

Análisis del Coste del Protocolo PIM-DM en topologías sin bucles

G. Maciá, J.E. Díaz-Verdejo
Departamento de Teoría de la Señal, Telemática y Comunicaciones. Universidad de Granada
ETSI de Ing. Informática. C/ Daniel Saucedo Aranda S/N
18071 - Granada
Teléfono: 958 24 23 04 Fax: 958 24 08 31
E-mail: jedv@ugr.es

Abstract *This work presents an approach to estimate the number of overhead packets, both for data an control traffic, generated by the use of the PIM-DM protocol. A loop-free network topology and equal transmission speeds and propagation times for all the links in the network are assumed. Although restrictive at a first glance, the results show a good performance in simulated real networks when mean values for the link parameters are used. The expressions are deduced from the protocol functioning, overcoming limitations and approximations of previously published works.*

1. Introducción

Multicast es un estándar IETF [1] propuesto para numerosas aplicaciones, entre otras, la red experimental Mbone [2], operativa desde el año 1992. Teóricamente, esta tecnología aporta una ventaja cuando es comparada con la transmisión unicast en aplicaciones diseñadas para dar servicio a un gran número de clientes o receptores de información distribuidos en la red. Esto se debe al hecho de que únicamente se envía un paquete multicast sobre cada uno de los enlaces en el camino hasta los miembros del grupo multicast, sin importar el número de receptores que se hayan unido al mismo. La duplicación de paquetes para poder alcanzar a todos los receptores sólo se realiza cuando el camino hacia ellos se separa. Este mecanismo implica, a priori, una reducción del ancho de banda utilizado para el envío de paquetes de datos, lo cual hace que la transmisión multicast resulte ventajosa en comparación con la tecnología unicast.

Sin embargo, si consideramos otros escenarios, tales como pequeñas redes y/o con pocos receptores, la ventaja de multicast no es tan clara. Esto se debe, principalmente, al coste de construcción del *árbol multicast*, necesario para la operación del algoritmo. Este árbol está formado por todos los enlaces y nodos en el camino, libre de bucles, para llegar desde la fuente a todos los receptores. Por supuesto, el descubrimiento del árbol y su mantenimiento, ya que puede cambiar en el tiempo, implican un coste de sobrecarga adicional, en términos de paquetes que es necesario enviar. En consecuencia, al no estar dicho coste presente en una transmisión unicast, aunque se utilice mayor ancho de banda para transmitir los datos, éste se puede ver compensado por la ausencia de los costes de sobrecarga. Si encontramos escenarios en los que las diferencias entre los costes de transmisión de datos entre ambas técnicas no son grandes, unicast puede llegar a ser una mejor elección que multicast.

El problema que se pretende abordar se puede formular, en consecuencia, en los siguientes términos: dado un escenario, con unas características conocidas, en el que se va a realizar una transmisión a múltiples receptores, determinar si resulta más ventajoso utilizar técnicas de transmisión multicast o, por el contrario, es más eficiente utilizar unicast. Evidentemente, para poder tomar la decisión sobre el tipo de transmisión elegir, deberíamos estimar previamente el coste de ambas técnicas en dichas redes.

El principal objetivo de este trabajo consiste, en consecuencia, en presentar una estimación del coste asociado al uso del protocolo de transmisión multicast PIM-DM. Este protocolo, especificado en [3], está recomendado para situaciones en que los receptores se encuentran ubicados de forma agrupada y no muy dispersa en la red.

El resto del artículo se organiza como se describe a continuación. En primer lugar se presentan y discuten los trabajos previos relacionados con este estudio. En la Sección 3 se realiza un análisis del cálculo del coste del protocolo, bajo unas hipótesis que permitan abordar el problema, tanto para la transmisión de datos como para el coste de sobrecarga. La Sección 4 muestra cómo extender los resultados previos a entornos reales con topologías sin bucles. Finalmente se exponen las conclusiones y algunas líneas de trabajo futuro.

2. Antecedentes

Varios trabajos han considerado el estudio comparativo, en términos de eficiencia,

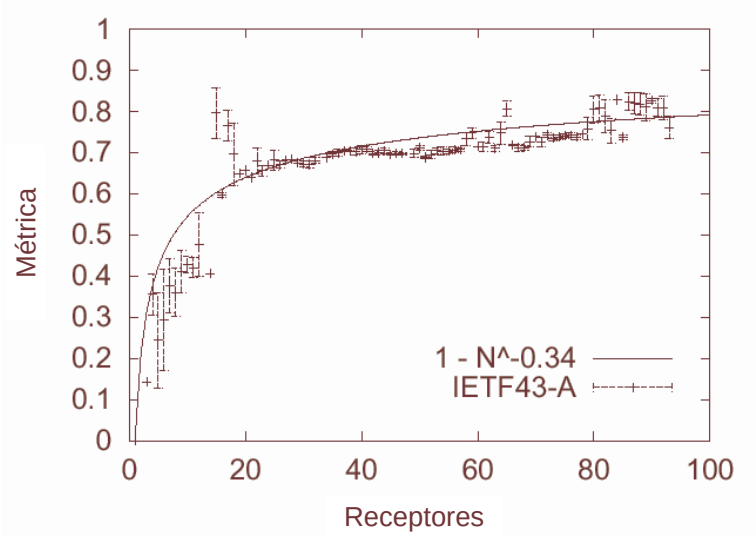


Figura 1: Ajuste experimental de la ley exponencial propuesta por Chuang y Sirbu. Tomado de [5].

de multicast y unicast, entre los que destacan [4] [5] [6] [7] [8].

Estos trabajos están basados en el estudio presentado por Chuang y Sirbu [4], en el que se establece una cuantificación del coste multicast en función del número de abonados a una transmisión en una red (número de receptores). Para ello se utilizan redes con topologías reales y generadas o simuladas, simultaneando numerosos tipos de topologías y tamaños de redes. Se postula una ley exponencial, ratificada con resultados experimentales de simulación, de acuerdo a

$$\frac{L_m}{L_u} = N^{0,8} \tag{1}$$

donde L_m es el tamaño del árbol multicast, en número de enlaces, L_u la longitud media del camino unicast hasta los receptores y N el número de receptores existentes en la red. Esta expresión fue validada experimentalmente (Fig. 1) por [5]. Como se puede apreciar en la figura, cuando el número de receptores es elevado, la aproximación es bastante ajustada. Sin embargo, esto no es cierto en el caso de que el número de receptores sea bajo. Adicionalmente, [7] acota aún más la validez de la Ec. (1), ya que demuestra que es válida para N no muy grandes, indicando que es necesario un tratamiento diferente para un número de receptores muy elevados.

Por otra parte, [4] propone un método analítico de cálculo para la estimación del coste de la transmisión multicast de acuerdo al protocolo PIM-DM. Para ello, el tráfico generado se clasifica en dos tipos conceptualmente diferentes: el *tráfico de datos* y el trafico asociado de *sobrecarga*. El *tráfico de datos* corresponde a paquetes de información que alcanzan algún sistema final receptor, es decir, algún miembro del grupo multicast. Por el contrario, el *tráfico de sobrecarga* está compuesto por los paquetes que se generan debido a las necesidades del protocolo (creación y mantenimiento del árbol multicast) y los que no alcanzan ningún receptor.

El coste del protocolo PIM-DM, C_{DM}, se puede expresar, en consecuencia, como suma de dos términos diferentes: el *coste del tráfico de datos*, CD_{DM}, y el *coste de sobrecarga*, CO_{DM}.

$$C_{DM} = CD_{DM} + CO_{DM} \tag{2}$$

La expresión propuesta para evaluar el coste de los paquetes de datos, (CD_{DM}), aunque originariamente descrita en términos de la tasa de tráfico que genera, puede transformarse de forma simple para que evalúe el coste de la transmisión de datos en función del número de paquetes de datos totales que atraviesan el árbol multicast en un tiempo T_{transm}, resultando [4]:

$$CD_{DM} = \alpha \cdot L_m \tag{3}$$

donde α es el número de paquetes transmitidos por la fuente y L_m el número de enlaces del árbol multicast.

Por otra parte, en este trabajo se presenta una expresión para el coste del tráfico de sobrecarga, CO_{DM}, que evalúa el número de paquetes generados por el mecanismo de *inundación y poda* empleado en el protocolo PIM-DM para la creación y mantenimiento del árbol multicast durante T_{transm}:

$$CO_{DM} = 2 \cdot (L'_m - L_m) \cdot F \tag{4}$$

donde L'_m es la longitud del *árbol de broadcast*, es decir, el número total de enlaces de la red considerada, y F representa el número de inundaciones durante el periodo de observación. Si consideramos un tiempo T_{transm}, el número de inundaciones puede expresarse, de acuerdo a las especificaciones del protocolo, de la forma:

$$F = \left\lceil \frac{T_{transm}}{\tau_{dm}} \right\rceil \tag{5}$$

siendo τ_{dm} el tiempo entre inundaciones periódicas, también denominado *temporizador de poda*, y donde $\lceil \cdot \rceil$ representa el operador entero superior.

Las expresiones proporcionadas en [4] constituyen una buena aproximación para estimar el coste en redes grandes y, principalmente, en escenarios en los que el coste de los datos es mucho mayor que el de sobrecarga. Ahora bien, cuando se consideran redes que no cumplen estas condiciones, dos aspectos importantes del comportamiento del protocolo PIM-DM, que no son considerados en la expresión que calcula el coste de sobrecarga, comienzan a tener relevancia y, por tanto, provocan desviaciones respecto del comportamiento real que hacen que dichas expresiones no sean aplicables en estos casos. Estas limitaciones no son mencionadas en [4], que supone la aplicabilidad de las expresiones indicadas en todos los casos.

Dos son los aspectos relevantes no considerados en la deducción de las expresiones del coste. En primer lugar, en la Ec. (4) se considera que, sobre todos los enlaces que no pertenecen al árbol multicast, únicamente se transmite, en cada inundación, un paquete de datos multicast y una respuesta de poda. Obviamente, esta consideración implica que el número de paquetes de sobrecarga corresponde al número de enlaces multiplicado por dos. Sin embargo, cuando un encaminador empieza a enviar paquetes de datos, no deja de hacerlo hasta recibir un mensaje de poda. Esto significa que debemos considerar que un encaminador, enviando a una tasa constante, podría mandar más de un paquete hasta recibir la respuesta de poda. Además, cada paquete enviado a través de un enlace que no pertenezca al árbol multicast originará su correspondiente respuesta de poda por parte del encaminador en el otro extremo. Por ello, la Ec. (4) se deberá modificar de modo que recoja este efecto.

En segundo lugar, podemos apreciar también que en la Ec. (5) se supone que se produce una inundación cada τ_{dm} segundos. Esta aproximación no es adecuada debido a que implica que el temporizador de poda se inicia con la recepción del primer mensaje de poda y que los siguientes mensajes de poda no lo ponen a cero. Como se puede comprobar en las especificaciones del protocolo [3], este temporizador se debe reajustar con la llegada de cualquier mensaje de poda. Así, se hace necesaria una modificación de la Ec. (4) que recoja los efectos de este comportamiento.

En conclusión, aunque las expresiones proporcionadas por Chuang [4] constituyen una buena estimación en escenarios en los que el coste de sobrecarga no es representativo frente al de datos, es necesario modificar estas expresiones para poderlas aplicar en redes donde dicho coste se convierte en un factor relevante.

3. Estimación de coste para el protocolo PIM-DM

En la evaluación del coste para el protocolo PIM-DM debemos tener en cuenta, como se ha comentado en la sección anterior, dos costes separados: el de la transmisión de los datos, o *coste de datos*, CD_{DM}, y el originado por la generación y mantenimiento del árbol multicast, denominado *coste de sobrecarga*, CO_{DM}.

En primer lugar, la expresión para el coste de datos proporcionada por la Ec. (3) es correcta incluso en escenarios con redes pequeñas o donde el coste de sobrecarga sea significativo, debido a que el comportamiento del protocolo que no ha sido modelado en [4] no afecta a los enlaces pertenecientes al árbol multicast. Sin embargo, esto no es cierto respecto de la Ec. (4) que evalúa el coste de sobrecarga.

Para abordar más fácilmente el problema, dividiremos la estimación del coste de sobrecarga en dos términos: el coste debido a la generación de mensajes de poda, que denominaremos *coste de poda* (CO_{DM}^{poda}), y el coste adicional debido a los datos que circulan por enlaces no pertenecientes al árbol multicast, que es provocado por las inundaciones periódicas asociadas al protocolo PIM-DM, y que llamaremos *coste CBR* (CO_{DM}^{cbr}).

$$CO_{DM} = CO_{DM}^{poda} + CO_{DM}^{cbr} \tag{6}$$

Para evaluar estos términos, estableceremos las siguientes hipótesis de trabajo respecto del escenario a considerar:

- Existe una única fuente multicast transmitiendo paquetes de forma continua durante un tiempo T_{transm}.
- No existen fenómenos de pérdida de paquetes ni congestión en ninguna parte de la red.
- El tiempo de propagación y las velocidades de transmisión son iguales en todos los enlaces de la red. Esta suposición, aunque no es muy realista, se realiza con la finalidad de simplificar las ecuaciones. En la Sección 4 se discutirán algunos aspectos relativos a la aplicabilidad de las expresiones obtenidas a entornos reales.

3.1. Estimación del coste de poda para ramas de n-saltos

Para estimar el coste de poda comenzaremos por un modelo simple para, posteriormente, extenderlo a escenarios más complejos. Consideraremos, inicialmente, una red como la representada en la Fig. 2, en la que existe un único enlace. En consecuencia, existirán dos encaminadores unidos por un enlace punto a punto. Uno de ellos (nodo 1) actúa como fuente multicast y el otro (nodo 2) se

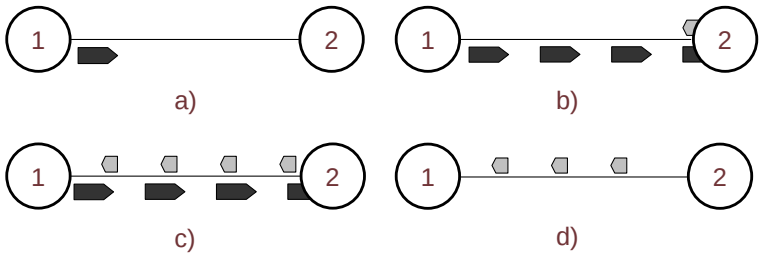


Figura 2: Mecanismo de inundación y poda en una rama de un salto. Fases 1 (a), 2 (b), 3 (c) y 4 (d).

supondrá un elemento de red que no pertenece al grupo multicast. El mecanismo de inundación y poda especificado en el RFC 1112 [1] se comporta de la siguiente forma. En t=0 (Fig. 2(a)), la fuente comienza a enviar paquetes. Cuando el primer paquete de datos llega al nodo 2 (Fig. 2(b)), éste genera una respuesta de poda a dicho mensaje. Dicha respuesta, al llegar al nodo 1 (Fig. 2(c)) provoca que la fuente deje de emitir mensajes. La fuente iniciará el temporizador de poda, τ_{dm} con la llegada del último mensaje de poda (Fig. 2(d)). Por supuesto, cada paquete de datos posterior que llega al nodo 2 genera el mismo comportamiento.

Sea t_1 el tiempo total necesario para transmitir un paquete de datos entre dos nodos vecinos y t_2 el empleado por la respuesta generada en el camino inverso. Estos dos tiempos serán diferentes, ya que las longitudes de los paquetes de datos y de respuesta lo son. Si denominamos t_0 la suma de ambos tiempos, es decir, al tiempo transcurrido desde que se inicia la transmisión de un paquete de datos hasta que se recibe la respuesta generada por su vecino, tendremos:

$$t_0 = t_1 + t_2 = 2 \cdot t_p + \frac{L_{cbr} + L_{prune}}{V_t} \qquad (7)$$

donde t_p es el tiempo de propagación sobre el enlace, L_{cbr} y L_{prune} son, respectivamente, los tamaños de los paquetes CBR y poda y, finalmente, V_t es la velocidad de transmisión sobre el enlace. Obviamente, esta expresión no considera aspectos como retrasos en colas o procesamientos en los encaminadores.

A partir del funcionamiento del protocolo y examinando la Fig. 2, podemos deducir que el número de mensajes de poda generados es igual al número de mensajes CBR enviados por la fuente. Así, definimos la fracción de paquetes enviados por la fuente, N, como:

$$N = \frac{t_0}{\tau_{cbr}} \qquad (8)$$

donde τ_{cbr} es el tiempo entre dos envíos consecutivos de paquetes de datos por la fuente. Evidentemente, N no tiene porqué ser un valor un entero, por lo que, en consecuencia, el número de paquetes generados durante F_1 inundaciones realizadas en un periodo de observación T_{transm} se obtiene a partir de:

$$CO_{DM}^{prune} = F_1 \cdot \lceil N \rceil \qquad (9)$$

Los resultados obtenidos para un único salto se pueden extender a ramas de n saltos. En primer lugar, consideraremos una extensión a 2 saltos para, posteriormente, aplicar los resultados a topologías en estrella con ramas de hasta n saltos en las que no existan bifurcaciones.

La Fig. 3 muestra gráficamente el mecanismo inundación y poda en una rama de dos saltos. El comportamiento es muy similar al del escenario anteriormente descrito: En la fase 1, el nodo 1 comienza a enviar paquetes de datos (Fig. 3.a); cuando los paquetes alcanzan el nodo 2 (Fase 2), éste los retransmite hacia el nodo 3 (Fig. 3.b); la llegada de paquetes al nodo 3 (Fase 3) provoca la generación de respuestas de poda (Fig. 3.c); y, finalmente, la llegada del primer mensaje de poda al nodo 2 provoca que éste no emita más (Fase 4) y que, a su vez, genere un mensaje de poda hacia el nodo 1 (Fig. 3.d), que cuando llega a su destino, hace que la fuente deje de emitir e inicie el temporizador de poda τ_{dm}.

Observamos en este escenario que se generan mensajes de poda solamente en dos casos: como respuesta a un paquete de datos y con la llegada, en un nodo intermedio (nodo 2), del primer mensaje de poda. Este segundo caso solo produce un paquete en cada inundación del protocolo por cada nodo intermedio, mientras que el primer caso produce $\lceil N \rceil$ paquetes en cada enlace. Por tanto, en el caso de 2 saltos, podemos estimar el coste de sobrecarga de acuerdo a

$$CO_{DM}^{prune} = F_2 \cdot \lceil 2 \cdot N + 1 \rceil \qquad (10)$$

donde F_2 es el número de inundaciones que se producen en una rama con dos saltos.

En general, siguiendo los razonamientos anteriores, para una rama con n saltos tendríamos

$$CO_{DM}^{prune} = F_n \cdot \lceil n \cdot N + n - 1 \rceil \qquad (11)$$

Finalmente, si consideramos una topología en estrella con ramas de 1 a n saltos sin bifurcaciones, podemos concluir que, si existen L_i ramas de i saltos, el número de mensajes de poda generados es:

$$CO_{DM}^{prune} = \sum_{i=1}^{n} L_i \cdot F_i \cdot \lceil i \cdot N + i - 1 \rceil \qquad (12)$$

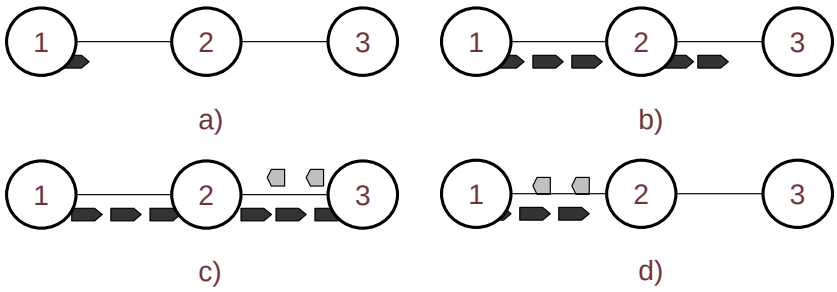


Figura 3: Mecanismo de inundación y poda en una rama de dos saltos. Fases 1(a), 2(b), 3(c) y 4(d).

donde F_i es el número de inundaciones que tienen lugar en una rama de i saltos. Observamos que, bajo las hipótesis presentadas, el intervalo hasta la llegada del último mensaje de poda al nodo fuente de la transmisión depende únicamente del número de saltos de la rama. Es por ello que podemos asumir, sin pérdida de generalidad, que el número de inundaciones será el mismo en todas las ramas de igual longitud en número de saltos.

3.2. Estimación del coste CBR para ramas de n-saltos

El *coste CBR*, como se ha indicado anteriormente, está asociado a los paquetes de datos que circulan por la red y no alcanzan ningún receptor, es decir, viajan por ramas que no pertenecen al árbol multicast. Estos paquetes son generados durante las inundaciones periódicas asociadas al protocolo PIM-DM, por lo que se producirán hasta que las ramas correspondientes sean convenientemente podadas tras recibirse los mensajes de poda.

Se puede comprobar en el escenario de un salto mostrado en la Fig. 2 que el número de paquetes generados por la fuente es igual al número de mensajes de poda producidos por el nodo 2. Por tanto, para este escenario, a partir de la expresión (9) podemos concluir que:

$$CO_{DM}^{cbr} = CO_{DM}^{prune} = F_1 \cdot \lceil N \rceil \qquad (13)$$

En el escenario con una rama de dos saltos mostrado en la (Fig. 3) se pueden establecer varias fases diferenciadas por el número de paquetes CBR generados en la rama completa. Así, existen dos periodos de tiempo en los que únicamente hay un paquete de datos CBR en la rama completa por cada uno de los que genere la fuente. Estos periodos corresponden a la situación inicial, en la que únicamente se está transmitiendo por el primero de los enlaces, es decir, el intervalo $(0,\ t_1)$ -Fig. 3(a)-, y a la situación en la que el segundo enlace ha sido ya podado, lo que corresponde al intervalo $(2t_1 + t_2,\ 2t_0)$ -Fig. 3(d)-. Sin embargo, en el intervalo $(t_1,\ 2t_1 + t_2)$ -Fig. 3(b) y 3(c)- existen dos paquetes por cada uno que envía la fuente, ya que ambos enlaces se encuentran activos. Por tanto, el número de paquetes CBR será igual al número de paquetes de poda durante $t_1 + t_2 = t_0$ y dos veces este número durante el mismo tiempo t_0. De esta forma, el coste en una rama de dos saltos será:

$$CO_{DM}^{cbr} = F_2 \cdot (\lceil N \rceil + \lceil 2 \cdot N \rceil) \qquad (14)$$

Extendiendo el razonamiento a una rama con n saltos, se puede deducir la expresión del coste de sobrecarga para una topología en estrella genérica con profundidad máxima n, compuesta por L_i ramas de i saltos, resultando:

$$CO_{DM}^{cbr} = \sum_{i=1}^{n} F_i \cdot L_i \cdot \sum_{j=1}^{i} \lceil i \cdot N \rceil \qquad (15)$$

3.3. Número de inundaciones

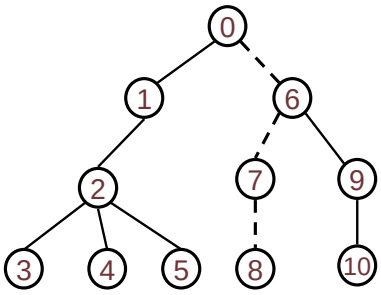
Como se ha indicado en la Sección 3.1, el número de inundaciones en una rama depende exclusivamente de su profundidad, es decir, del número de saltos en dicha rama. Supongamos en lo que sigue una rama de longitud i. Para evaluar el número de inundaciones que tendrán lugar hemos de revisar la operación del protocolo PIM-DM y los *temporizadores de poda* asociados.

De acuerdo a la especificación del protocolo, cada vez que se recibe un mensaje de poda se reinician los temporizadores de poda. Por tanto, es necesario estimar el tiempo T_i que transcurre entre el envío del primer paquete de datos sobre una rama de i saltos que deba ser podada (paquetes CBR) y la recepción del último mensaje de poda procedente de dicha rama. Este tiempo se puede dividir en dos contribuciones: *a)* el tiempo entre el envío del primer mensaje CBR y la llegada del primer mensaje de poda y *b)* el tiempo entre la llegada del primer y el último mensaje de poda.

El máximo tiempo entre el envío del primer paquete CBR y la recepción del primer mensaje de poda es i veces t_0, ya que el paquete CBR deberá llegar al nodo más profundo de la rama y su respuesta de poda debe volver a la fuente. Por otro lado, el máximo tiempo entre la recepción del primer y el último mensaje de poda es t_0, debido a que se puede generar un paquete CBR justo antes de recibir la respuesta de poda.

Por tanto, podemos concluir que:

$$T_i = \begin{cases} (i+1) \cdot t_0 + \tau_{dm} & \text{si } \tau_{cbr} < t_o \\ i \cdot t_0 + \tau_{dm} & \text{si } \tau_{cbr} \geq t_o \end{cases} \qquad (16)$$



a)

Valores de los coeficientes
$L_m = 3$
$L_3 = 1$
$L_2 = 1$
$L_1 = 2$

b)

Figura 4: Ejemplo de topología en árbol a la que es posible extender las expresiones deducidas. La fuente se situaría en la cabecera del árbol. a) Topología de ejemplo. En línea discontinua se muestra el árbol multicast considerado; b) Valores de los coeficientes.

donde hemos tenido en cuenta la emisión completa de un paquete en el caso de que se reciba el mensaje de poda tras el inicio de la misma.

Finalmente, el número de inundaciones que se producen en un intervalo T_{transm} sobre una rama de produndidad i será:

$$F_i = \left\lceil \frac{T_{transm}}{T_i} \right\rceil \tag{17}$$

3.4. Extensión a topologías en árbol

Las expresiones deducidas en (12) y (15) permiten evaluar los costes en topologías en estrella, esto es, compuestas por ramas simples sin bifurcaciones. Se puede extender la aplicabilidad de dichas expresiones a topologías en árbol (Fig. 4), esto es, topologías sin bucles, sin más que realizar un recuento del número de ramas existentes con una profundidad dada. Es decir, de acuerdo a la terminología utilizada, únicamente es necesario obtener los valores de los coeficientes L_i.

Dado un número de saltos, i, el número de ramas con dicho número de saltos sin bifurcaciones existentes en una topología en árbol cualquiera con profundidad máxima n puede determinarse explorando la red mediante el siguiente algoritmo:

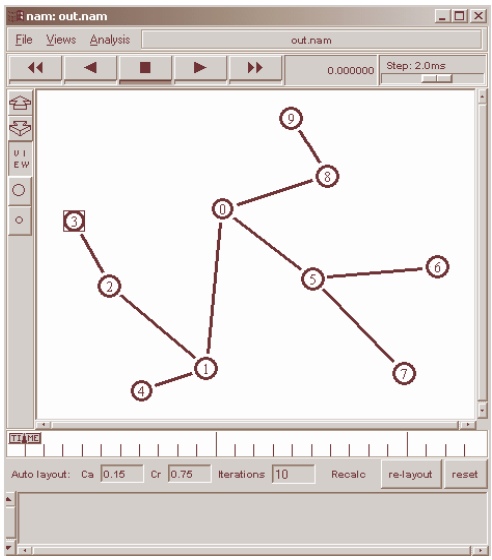
```
Hacer L_i = 0 ∀i/1 ≤ i ≤ n
Para cada rama que no pertenezca al
  árbol multicast
{
   1.1 Tomar el camino más largo P (en
     número de saltos) desde la fuente.
   1.2 Incrementar L_i en una unidad,
     siendo i el número de saltos del
     camino P
   1.3 Para cada nodo en el camino P,
     ir a 1.1.
}
```

4. Estimación de costes en redes sin bucles

Las expresiones (12), (15) y (17), junto con el algoritmo para evaluar los coeficientes L_i, han sido evaluadas experimentalmente usando el simulador NS2 [9]. Para ello se ha generado un conjunto de 36 topologías en árbol al azar de acuerdo a las hipótesis establecidas en la Sección 3, mediante el uso del generador de topologías GT-ITM [9]. Los valores obtenidos a partir de la simulación concuerdan con los previstos a partir de las expresiones mencionadas. A modo de ejemplo, la Fig. 5 muestra uno de los escenarios evaluados y los valores obtenidos.

De acuerdo a las condiciones establecidas en la Sección 3, la estimación de costes ha sido obtenida bajo la hipótesis de que todos los enlaces en la red presentan la misma velocidad de transmisión y tiempo de propagación. Evidentemente, las redes reales no cumplen estos requerimientos, por lo que se han realizado varios experimentos con el objetivo de obtener una estimación sobre la validez de la aproximación.

La aplicación a redes reales sin bucles se realizaría utilizando en las expresiones los valores medios de la velocidad de transmisión y del tiempo de propagación para todos los enlaces de la red. De esta forma, se han llevado a cabo simulaciones experimentales usando topologías con un amplio rango de valores para el ancho de banda y tiempo de propagación. Los resultados de todas las simulaciones presentan una mayor desviación en la estimación del coste de sobrecarga (variaciones sobre el 6.5 % con varianzas de ±7 %), que en la estimación del coste de poda (variaciones medias del 1.9 %, con bajas varianzas ±2,5 %). En todo caso, las expresiones propuestas aportan una estimación mucho más adecuada que las presentadas en [4], como se muestra en la Fig. 6, ya que aquéllas no modelan correctamente el cálculo de los paquetes de sobrecarga del protocolo. Ello hace que, cuando dicho término es significativo, es decir, cuando el número de receptores no es muy elevado o cuando



a)

Datos de simulación	
Duración	2,8 s
Velocidad enlaces	100 Mbps
Tiempo propagación enlaces	0,16 s
Velocidad fuente	1000 paq./s

Resultados obtenidos		
	Reales	Calculados
Número paquetes CBR	4489	4489
Número paquetes de poda	2289	2289

b)

Figura 5: Ejemplo de red evaluada: a) Topología en árbol seleccionada, b) Características de la red y resultados obtenidos.

el tiempo de ida y vuelta es mayor que el tiempo entre emisiones de la fuente, se observa que la estimación propuesta en [4] es deficitaria. Las expresiones deducidas en este trabajo se adaptan mejor a este tipo de condiciones.

La variación de los resultados de unas topologías a otras depende de factores tales como la propia varianza de los tiempos de propagación o la distribución topológica de los mismos en cuanto a cercanía a la fuente. Tener en cuenta estos valores implicará adoptar valores diferentes a la media en la expresión de cálculo, a fin de reducir las desviaciones. Aunque éste es un aspecto que queda fuera del alcance de este trabajo, se observa que las desviaciones permiten tener un margen de error suficientemente estrecho para decidir si es mejor la transmisión multicast o unicast.

En cuanto a la velocidad de transmisión, se puede observar que la sensibilidad de la expresión es muy baja con respecto a esta variable, siempre que se cumpla la condición realista de tener velocidades que hagan que el tiempo de transmisión de los paquetes sea despreciable respecto al tiempo de propagación de los mismos:

$$t_0 = t_1 + t_2 = 2 \cdot t_p + \frac{L_{cbr} + L_{prune}}{V_t} \approx 2 \cdot t_p \quad (18)$$

Como conclusión de este análisis, observamos que el paso de topologías experimentales e ideales, en las cuales los anchos de banda y tiempos de propagación son iguales para todos los enlaces, a topologías con valores realistas de estas variables nos permite tener un margen de error suficientemente estrecho, definido por la variación del tiempo de propagación fundamentalmente. La elección de un criterio adecuado para el valor de esta variable nos permitirá acotar en cierto modo el margen de error obtenido y, por tanto, afinar aún más para tomar la decisión de realizar una transmisión multicast o unicast en un entorno real.

5. Conclusiones y trabajo futuro

En este trabajo se han mejorado las expresiones propuestas en [4] para la evaluación del coste del protocolo PIM-DM, en términos del número de paquetes. Dado que la propuesta de Chuang no es completamente adecuada para su aplicación en escenarios donde el coste de sobrecarga es relevante, por no considerar algunos aspectos de PIM-DM, se han incorporado estos fenómenos obteniendo expresiones que evalúan este coste en topologías sin bucles. Aunque las nuevas expresiones se han deducido con la condición de tener velocidades de transmisión y retardos en los enlaces de valores homogéneos, su aplicación en redes reales sin bucles muestra una clara mejoría en comparación con los resultados que aporta [4].

Los trabajos futuros deben ser dedicados a mejorar la aplicabilidad de las expresiones a redes reales incluyendo modelos detallados para el tiempo de propagación y la velocidad de transmisión de los enlaces. En cualquier caso, la principal ventaja de las expresiones propuestas reside en su simplicidad, que permite una evaluación rápida y fácil. Por otro lado, un objetivo importante que debe abordarse como trabajo futuro es la extensión de las expresiones a topologías genéricas que no tengan que cumplir la condición de no tener bucles. Además, sería necesario la realización de este estudio para otros protocolos de transmisión multicast.

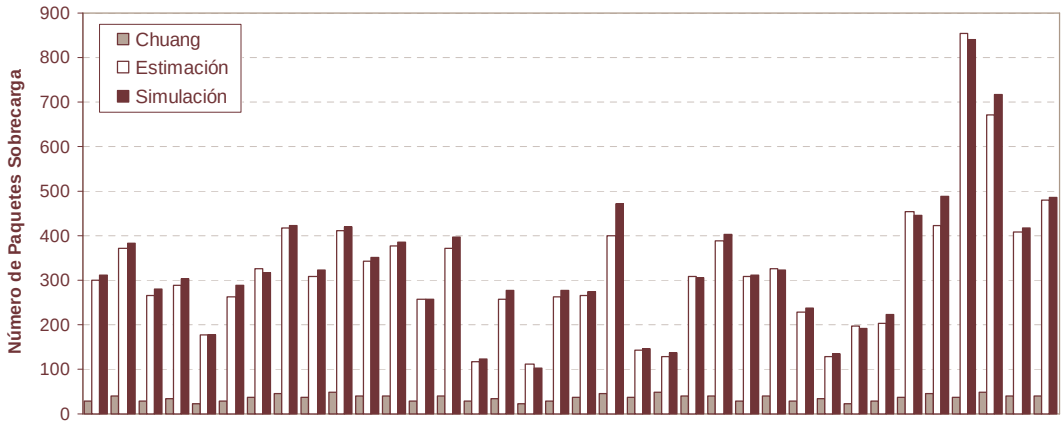


Figura 6: Número de paquetes de sobrecarga para varias topologías aleatorias, estimado según [4] (serie *Chuang*), según las expresiones propuestas (serie *Estimación*) y obtenido por simulación usando *Network Simulator 2* (serie *Simulación*).

Agradecimientos

Este trabajo ha sido parcialmente subvencionado por el MCYT a través del proyecto SERVIRA (TIC2002-02798, 70 % fondos FEDER).

Referencias

[1] S. Deering, *Host Extensions for IP Multicasting*, RFC 1112 (1989)

[2] H. Eriksson. *Mbone: the multicast backbone.* Communications of the ACM 37(8), pp 54-60 (1994).

[3] S. Deering, D. Estrin, D. Farinacci, V. Jacobson, A. Helmy, D. Meyer, L. Wei. *Protocol Independent multicast version 2 Dense mode specification.* Internet Draft, 1999.

[4] J. Chuang and M. Sirbu. *Pricing multicast communication: A cost-based approach*, Anual Conference of the Internet Society (INET'98), pp. 281-297.

[5] R.C.Chalmers and K.C. Almeroth. *Modeling the branching characteristics and efficiency gains in global multicast trees*, in IEEE INFOCOM, Apr. 2001, pp. 449-458.

[6] G.Phillips, S. Shenker, H. Tangmunarunkit. *Scaling of multicast trees: Comments on the Chuang-Sirbu scaling law*, in Proceedings of INET'98, pp. 41-51, Geneva (Suiza), 1998.

[7] P.V.Mieghem, G. Hooghiemstra and R. van der Hofstad. *On the Efficiency of multicast*, IEE/ACM TRANSACTIONS ON NETWORKING, nº 6, pp. 719-732 (2001).

[8] K.Calvert, M. Doar and E. Zegura. *Modeling Internet Topology*, IEEE Transactions on Communications, pages 160-163, 1997.

[9] http://www.isi.edu/nsnam/ns

Arquitectura de red para la automatización de pruebas

Alfredo Beaumont, José Oscar Fajardo, Eva Ibarrola, Cristina Perfecto
Networking, Quality and Security Research Group.
Departamento de Electrónica y Telecomunicaciones
Euskal Herriko Unibertsitatea/Universidad del País Vasco
ETSI de Bilbao. C/ Alda. Urquijo S/N. 48013 - Bilbao (Bizkaia)
Teléfono: 94 601 73 08. Fax: 94 601 42 59
E-mail: {jtbbesaaljtbfapoj}@aintel.bi.ehu.es, {jtpibaraljtppeamc}@bi.ehu.es

Abstract *One of the most important tasks of a researcher consists in performing tests to validate the results of the research done. In a network related research work, with complex infrastructure and many configuration parameters, testing our work is usually rather mechanic, tedious and error prone if tests are done manually. In the Networking, Quality and Security Research Group, we have developed a generic test framework to automate this important task, composed of four kind of logical elements: a manager, agents, daemons and formatters. With these four elements, every phase of the process is automated, from test configuration to result formatting. The deployment of this framework has drastically reduced the time needed for the test phase and the number of errors due to mistakes. It has also allowed us to share resources among projects more easily and to schedule test sets when the devices are not being used.*

1. Introducción

La validación de hipótesis es una de las tareas más importantes en el mundo de la investigación. Los resultados de las pruebas deben ser claros y no deben dejar duda alguna sobre la viabilidad de la hipótesis planteada. Muchas veces hasta que se obtienen los resultados deseados es necesario repetir varias veces las pruebas debido a errores humanos o a hipótesis de partida erróneas. En el mundo de la telemática, los resultados de la investigación están inherentemente unidos a una correcta validación mediante el desarrollo de prototipos que implementen las funcionalidades deseadas. Estas pruebas requieren la mayor parte de las veces de una serie de dispositivos conectados mediante algún tipo de red. Las configuraciones de estas redes suelen ser complejas y extremadamente variadas, por lo que normalmente las pruebas de validación se configuran de forma manual. Por ejemplo, en el caso de que se desee medir la eficiencia de varias tarjetas de red ante diferentes patrones de tráfico, diferentes velocidades y diferentes tamaños de paquete, sería necesario configurar individualmente cada uno de los elementos que forman parte de las pruebas. Además, las pruebas deberán repetirse para cada una de las tarjetas de red que se quieran testear. Esto acarrea una serie de problemas:

- La configuración de cada prueba requiere un tiempo considerable.
- La configuración de pruebas diferentes requiere la reconfiguración de algunos parámetros, lo que impide la realización de pruebas en lotes.
- La repetición de baterías de pruebas se hace dificultosa, ya que cualquier error o despiste en la configuración de algún parámetro puede provocar variaciones en los resultados.

Dentro de nuestro grupo de investigación Networking, Quality and Security nos hemos encontrado con esta problemática en diversos proyectos. La mayoría de las líneas de investigación relacionadas con el cálculo de parámetros de tráfico de red y el análisis de rendimiento en redes de alta velocidad necesitan la evaluación y validación de las hipótesis y de los desarrollos. La ejecución de multitud de pruebas sobre prototipos se ha venido desarrollando tradicionalmente de forma manual, con los problemas descritos anteriormente. Por ello, se consideró fundamental el diseño de una arquitectura que permitiera automatizar las tareas de configuración, ejecución y obtención de resultados de pruebas de validación. El objetivo fundamental es la obtención de resultados fiables de forma eficiente. Pero no menos importante es la generalidad de la arquitectura, es decir, esta arquitectura ha sido cuidadosamente diseñada para poder ser utilizada en multitud de entornos de red sin tener que realizar cambios en la estructura de los programas. Las especificaciones más relevantes que cumple esta arquitectura de pruebas son las siguientes:

- La configuración de los diferentes dispositivos que deben actuar en la prueba debe ser totalmente automatizable.
- Deben poder hacerse lotes de pruebas.
- Las baterías de pruebas deben ser reproducibles de forma automática.

- El sistema ha de ser resistente frente a errores, es decir, si se produce un error en una de las pruebas, debe afectar sólo a esa prueba y no al resto de las pruebas del lote.

- La arquitectura debe ser flexible y fácilmente ampliable, de tal forma que permita incorporar un nuevo dispositivo al entorno de pruebas de forma sencilla, aunque sea un tipo de dispositivo totalmente diferente al resto.

En este artículo se presenta la arquitectura genérica de pruebas que se ha desarrollado cumpliendo estos objetivos. El resto del artículo se organiza de la siguiente forma. En la sección 2 se describe la arquitectura general del diseño del sistema. En la sección 3 se describe el funcionamiento de la plataforma. En la sección 4 se expone un caso de uso, donde se explica la implementación realizada del sistema en un entorno de pruebas real y, por último, en la sección 5 se presentan algunas conclusiones con la experiencia obtenida y posibles líneas futuras de mejora del sistema.

2. Arquitectura

La arquitectura del sistema está compuesta por cuatro tipos de elementos lógicos. Cada uno de ellos implementan una funcionalidad perfectamente definida:

- Gestor. Es la interfaz con el usuario. Se encarga de gestionar todos los elementos de la red en función de la configuración que recibe del usuario, y también de almacenar los resultados obtenidos por estos elementos al finalizar cada prueba.

- Agentes. Se encargan de atender las peticiones del gestor y actuar sobre los diferentes dispositivos. Los agentes están siempre a la escucha y se encargan de arrancar y parar los demonios, y de recoger los resultados de su ejecución.

- Demonios. Son los encargados de actuar sobre los diferentes elementos físicos que participan en cada prueba. Su función puede ser muy variable, desde la configuración de algún elemento, hasta la ejecución de un programa, pasando por la obtención de información o estadísticas de algún elemento, etc.

- Los formateadores son los programas que se encargan de seleccionar y traducir la información almacenada por el gestor en formatos más apropiados para su representación.

Se puede observar la relación que existe entre cada uno de los elementos en la Fig. 1.

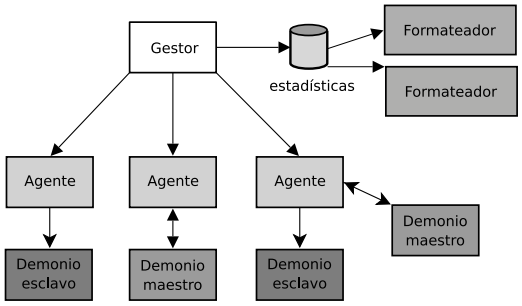


Figura 1. Esquema de la arquitectura de pruebas

A continuación se describe cada uno de los elementos de forma más detallada.

2.1. Gestor

El gestor proporciona la interfaz entre el entorno de pruebas y su administrador. Así pues, sus funciones principales consisten en interpretar la configuración de las pruebas, ejecutar las acciones indicadas por la configuración, y reunir los resultados de las mismas.

En la configuración se le indicarán el resto de elemenos que debe gestionar y las acciones que debe realizar el gestor. Sin embargo, con el objetivo de que la arquitectura sea lo más genérica posible, el gestor sólo debe interpretar la configuración genérica común a todos dispositivos, de tal forma que la adición de nuevos tipos de elementos a la arquitectura no implique la modificación del gestor.

Para realizar su tarea, el gestor debe comunicarse con uno o más agentes, indicándole a cada uno de los agentes las acciones a realizar, y comprobando los resultados que recibe del agente. Se trata pues de una arquitectura cliente/servidor, donde el gestor hace la función de cliente, conectándose a uno o varios servidores (agentes).

2.2. Agente

El agente se encarga de recibir los comandos del gestor para actuar sobre los demonios. Su función principal es la de atender los comandos del gestor e indicarle los resultados de sus acciones, incluyendo los posibles errores que se produzcan durante la ejecución de dichas acciones fortaleciendo la arquitectura frente a errores.

Un agente puede actuar sobre varios demonios de forma simultánea, por lo que presenta una arquitectura de servidor concurrente.

La comunicación entre el agente y los demonios está compuesta de tres fases:

1. Inicialización del demonio. El agente se encarga de arrancar el demonio.

2. Terminación del demonio. La terminación puede producirse por parte del agente o por parte del demonio, en función del tipo de demonio.

3. Recolección de estadísticas. El agente recoge las estadísticas que el demonio ha almacenado durante su ejecución.

Para realizar estas tareas es necesario definir un protocolo de comunicación entre el gestor y el agente mediante el cual se puedan invocar órdenes y recolectar información. Este protocolo es de nivel de aplicación y corre sobre TCP/IP. La Fig. 2 muestra un ejemplo del intercambio de información mediante este protocolo.

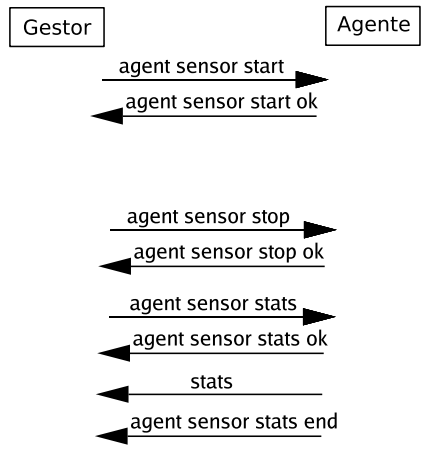


Figura 2. Intercambio de mensajes del protocolo

2.3. Demonios

Los demonios son los encargados de realizar acciones sobre los dispositivos. Estas acciones pueden ser de muy diversa índole, desde la modificación de una cola en el núcleo de un sistema operativo, hasta la recogida de estadísticas de uso de un switch, pasando por la ejecución de aplicaciones, etc. En definitiva, los demonios constituyen la parte específica de la arquitectura. Un demonio puede ejecutar cualquier acción, siempre que se ajuste a una interfaz predefinida para el intercambio de comandos y datos con el agente correspondiente. Así, un demonio puede desde ser un simple shell script que modifique ciertos parámetros del sistema, hasta una compleja aplicación de usuario. Esto viene determinado por la complejidad de la prueba, pero no está impuesto desde la arquitectura.

Dentro de los demonios se pueden diferenciar dos tipos, en función del tipo de relación que mantienen con el agente:

2.3.1. Maestros

Este tipo de demonios tienen cierta inteligencia. Será el agente quien los arranque, pero ellos indicaran cuando ha terminado su trabajo en función de lo que tarden en ejecutar la petición se ha sido realizada. Este tipo de demonios permite la realización de acciones asíncronas.

Hay que tener en consideración que la duración total de la prueba está supeditada al tiempo de ejecución de los maestros.

2.3.2. Esclavos

Este tipo de demonios se encargan de las tareas síncronas. Recibe la denominación de “esclavos” porque entre sus funciones no está la de determinar el fin de su ejecución. El agente controla tanto su arranque como su finalización. Su finalización la determinará el gestor en base al estado de los maestros.

2.4. Formateadores

Los formateadores son los encargados de tratar la información almacenada por el gestor tras una o varias pruebas. Hay dos aspectos principales que definen el comportamiento de un formateador:

1. Los datos sobre los que actúa. La cantidad de información generada por una prueba puede ser grande y de tipo muy variado. Es probable que interese extraer sólo parte de esa información para conocer el comportamiento de algún determinado parámetro.

2. El formato de salida. El objetivo final del formateador consiste en proporcionar la información tratada en un formato que pueda facilitar su interpretación o representación, bien haciendo uso de programas externos, proporcionando los formatos de hojas de cálculo o programas de gráficado, u obteniendo directamente formatos gráficos, tablas...

3. Funcionamiento

Una vez vista la arquitectura general del sistema, en esta sección se describe el funcionamiento del mismo, indicando los pasos de los que se compone el proceso de realización de una serie de pruebas y los elementos que intervienen en cada uno de estos pasos.

3.1. Configuración

En primer lugar ha de realizarse la configuración de la serie de pruebas que se quiere lanzar. Para la configuración de las pruebas se ha utilizado tecnología XML[1]. Consiste en un fichero con un conjunto de etiquetas y atributos XML. En cada fichero de configuración se puede definir una batería completa de pruebas, utilizando la etiqueta “test”. La configuración de cada una de las pruebas de la serie se realiza dentro de esta etiqueta. A continuación se muestra un ejemplo de un fichero de configuración de las pruebas.

```
<tests>
 <test name="example" comment="example test">
  <agent type="slave" name="sensor" addr="10.0.0.1"
   port="3333">
   <dev>eth0</dev>
   <rules>null.xml</rules>
   <proc>1</proc>
  </agent>
  <agent type="master" name="lambda" addr="10.0.0.2"
   port="5555">
   <num>10000</num>
```

```
    <rate>10000</rate>
    <src>192.168.1.91.3000</src>
    <dst>192.168.1.183.5000</dst>
    <dev>eth1</dev>
    <len>1500</len>
   </agent>
  </test>
</tests>
```

Hay que destacar la simplicidad con la que se pueden definir las pruebas. No obstante dada su naturaleza repetitiva, las pruebas normalmente suelen ser iguales a excepción de uno o dos parámetros que son los que se están testeando. Por ello se han desarrollado una serie de herramientas que facilitan su edición. Estas herramientas generan ficheros XML de pruebas a partir de unos ficheros de configuración, como se muestra en la Fig. 3. XSLT [2] facilita este trabajo.

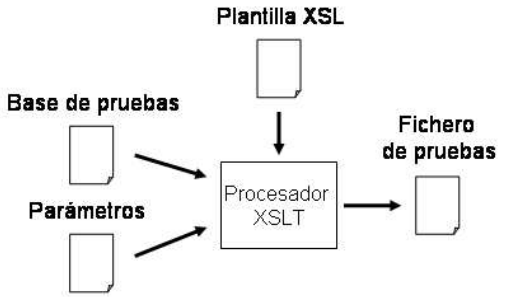


Figura 3. Generación de la configuración con XSLT

El fichero denominado "Base de pruebas" contiene información acerca de en que direcciones se encuentran los demonios y cuales son los valores de los parámetros por defecto. En el fichero "Parámetros" se indican qué valores deben cambiar en cada prueba y con qué incremento. Con estos datos junto a la plantilla XSL y el pocesador XSLT es capaz de generar ficheros de pruebas muy fácilmente.

El fichero "Base de pruebas" es muy similar al fichero de configuración mostrado anteriormente. Si añadimos un fichero de "Parámetros" como el que se muestra a continuación, obtendremos fichero de configuración de pruebas con una batería de 24 pruebas, en las que la tasa de inyección de paquetes aumentará de 5000 en 5000 paquetes por segundo

```
<?xml version="1.0" encoding="UTF-8"?>
<param>
 <name>rate</name>
  <initial_value>5000</initial_value>
  <increment>5000</increment>
  <times>24</times>
</param>
```

La primera acción del gestor es interpretar el fichero de pruebas generado y configurar cada uno de los elementos que tenga.

El orden de ejecución de cada prueba dentro de la serie es secuencial, por lo que pueden utilizar los mismos recursos o se les puede asignar el mismo nombre. Cada prueba sea una unidad de ejecución, en principio independiente del tiempo y/o orden en el que se ejecute.

Dentro de cada prueba, se define la localización de los agentes que participarán en la prueba, y la configuración de los demonios que se deberán ejecutar.

Cada demonio a ejecutar tendrá su correspondiente entrada dentro de la configuración de la prueba. La información asociada a cada demonio es la siguiente:

- El tipo de demonio, que puede ser maestro o esclavo.
- El nombre del demonio a ejecutar. Este nombre es lo que utilizará el agente, le permitirá conocer qué debe ejecutar.
- La dirección y puerto del agente. El gestor se conectará a esta dirección y puerto, que es donde debe estar escuchando el agente que gestionará el demonio.
- Los parámetros de ejecución del demonio. Además de lo anteriormente descrito, en el caso de que el demonio tuviera parámetros de configuración, estos también se incluirían aquí.

3.2. Inicialización de las pruebas

Una vez realizada la configuración de una o más pruebas, para que éstas se realicen hay que ejecutar el gestor. El gestor interpreta la configuración y determina los demonios que se han de utilizar y los agentes que los han de gestionar.

El gestor se conecta a todos los agentes, y de forma secuencial, envía un comando de arranque de todos los demonios a los agentes correspondientes, siguiendo el orden establecido en la configuración de la prueba. Para ello el gestor utiliza los elementos de configuración de cada demonio. De estos elementos, utiliza la dirección y el puerto del agente para conectarse a él y el tipo de demonio para definir la comunicación con el agente. Sin embargo, no evalúa ni el nombre del demonio ni sus parámetros, que recoge según el formato atributo-valor. De esta forma, la inclusión de un nuevo demonio no implica modificaciones en el gestor.

Los agentes, al recibir el comando de arranque por parte del gestor, arrancan los demonios con los parámentos indicados en el propio comando, en función del tipo de demonio del que se trate.

En el caso de un maestro, éste se ejecuta y realiza su acción, hasta que ésta finalice, momento en el que devolverá un mensaje al agente indicando el resultado de la operación, tanto si ha tenido éxito como si se ha producido un error.

En el caso de un esclavo, cuando el agente envía el comando de inicialización, si todo funciona correctamente, el demonio almacena un identificador único en un lugar conocido tanto por el demonio como por el agente, que le permitirá a este último detenerlo cuando llegue el momento. Además, indica al agente que ha arrancado correctamente y continúa realizando su trabajo en segundo plano. En el caso de que se produzca un error en el arranque del demonio, éste se indicará con un código de error al agente.

En el caso de que se haya producido un error y el agente lo haya detectado, éste se lo notificará al gestor. Si todo ha ido bien, el agente también devolverá un mensaje indicándolo.

Ni los agentes ni los demonios requieren acceder al fichero de configuración XML para realizar su función, la arquitectura es totalmente jerárquica.

Algunas de las posibles causas de los errores en la inicialización de las pruebas son las siguientes:

- El fichero de configuración es erróneo.
- Los agentes no están a la escucha o son inalcanzables. Es posible que no se hayan arrancado los agentes en todos los dispositivos, o que ocurran fallos en la red.
- Los demonios no se encuentran en el dispositivo correspondiente.
- Los parámetros de arranque del demonio son incorrectos. Hay que tener en cuenta que, por flexibilidad, ni el gestor ni el agente hacen comprobación alguna de estos parámetros, por lo que pueden ser no válidos.

Todos estos errores son detectables por el sistema.

3.3. Finalización de las pruebas

En el apartado anterior se ha comentado que los maestros no devuelven un mensaje al agente hasta que finaliza su trabajo manteniendo al agente a la espera, y éste, a su vez, al gestor.

Por tanto, hasta que no finalicen todos los maestros, el gestor se mantiene a la espera. Hay que tener en cuenta que el gestor debe ser capaz de gestionar el comportamiento de múltiples maestros de forma concurrente.

Cuando los maestros terminan el gestor continúa con las pruebas y comunica a los agentes que terminen la ejecución de los esclavos. Cada agente utiliza el identificador asociado a cada demonio para terminarlos de forma ordenada, y devuelve un mensaje con el resultado de la operación al gestor. Es importante tener en cuenta que el orden de parada de los demonios puede tener relevancia y por ello ha de realizarse de forma ordenada.

El caso más simple es aquél en el que se emplea un sólo maestro. Cuando éste termine su tarea, el agente terminará los esclavos y la prueba finalizará.

En el caso de varios maestros aparece la dificultad añadida de comprobar cuándo han finalizado todos ellos. El gestor debe de atenderlos de forma concurrente y no proceder a la finalización de la prueba hasta que no hayan devuelto todos el control.

3.4. Recogida de datos

Una vez finalizada una prueba, se procede a la fase de recogida de los datos asociados a la prueba. La obtención de resultados es una tarea asociada a cada demonio, de forma que ni el gestor ni los agentes requieren conocer o procesar estos resultados.

Los demonios, en el caso de que así lo requieran, almacenan el resultado de su ejecución en un fichero cuya localización es conocida tanto por el demonio como por el agente encargado de gestionarlo.

Cuando se termina la ejecución de los demonios, el gestor envía a todos los agentes un comando de petición de estadísticas. Los agentes recogen la información almacenada por los demonios y se la devuelven al gestor, que la almacenará en un fichero de resultados.

La información de las pruebas se almacena según un formato tipo, (unidad), valor. Es decir, en primer lugar se almacena el tipo de dato, en segundo lugar la unidad correspondiente en el caso de que se trate de algún tipo de medida, y en tercer lugar el valor del dato. Todas estas medidas se almacenan de forma secuencial, en "bruto", sin que el gestor las analice. Es por ello que se hace necesario un proceso de procesamiento posterior.

3.5. Presentación de resultados

Esta última fase consiste en el uso de formateadores sobre los resultados de las pruebas almacenados por el gestor, con el fin de seleccionar la información pertinente y representarla en un formato más adecuado.

4. Implantación en un entorno real

En esta sección se describe un entorno real de aplicación de la arquitectura de pruebas descrita en este artículo para demostrar la flexibilidad y la potencia de la arquitectura. Sin embargo, como ya se ha mencionado anteriormente, hay que tener en cuenta que dada la generalidad de la arquitectura su uso no se restringe exclusivamente a un entorno de análisis de tráfico como el que se va a describir a continuación. Esta arquitectura se ha utilizado para la automatización de pruebas de validación de un sensor multiprocesador dedicado al análisis pasivo de tráfico de red. Para esta aplicación en concreto las pruebas son fundamentales ya que es necesario probar el rendimiento del mismo cuando las características del tráfico que circula por la red cambian [3].

Las características más importantes del tráfico son la longitud del paquete, la distribución de la llegada de paquetes al sensor, la tasa de llegada,etc. Para cada uno de estos parámetros es necesario hacer pruebas repetidas veces para comprobar la estabilidad de los resultados.

La infraestructura, que se puede observar en la Fig. 4 se compone de una serie de máquinas interconectadas mediante elementos de red. Se pueden observar redes diferenciadas. Una de las redes es la que incluye e interconecta los elementos que deben participar en la prueba y la otra es la que comprende los elementos encargados de la gestión de las pruebas.

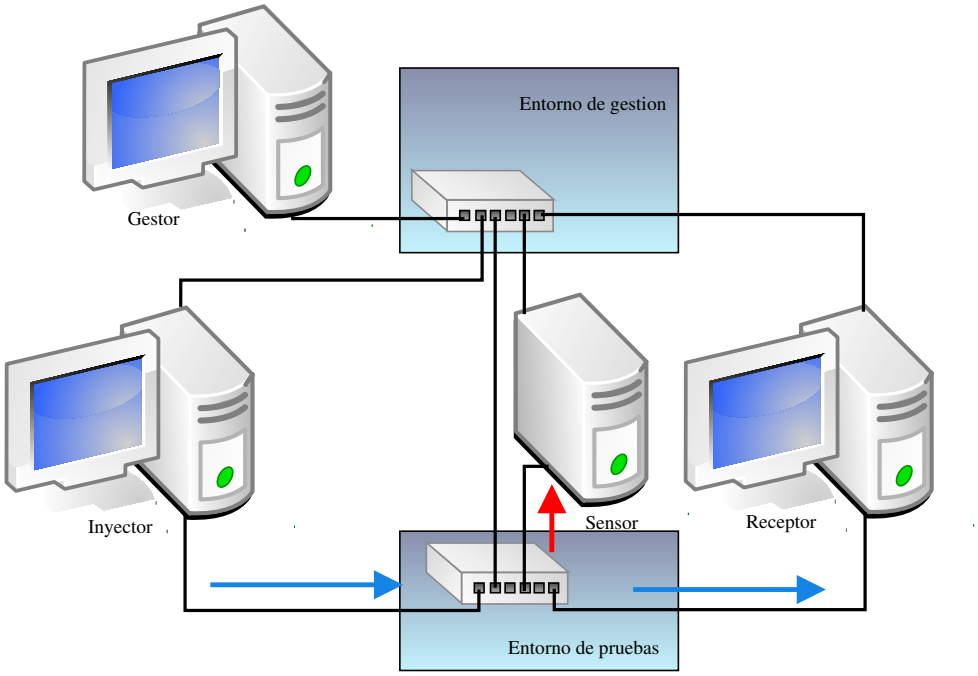


Figura 4. Esquema de la implantación de la arquitectura de pruebas

La utilización de dos redes separadas se hace necesaria para que el intercambio de información entre los elementos de gestión no alteren los propios resultados de las pruebas.

4.1. Entorno de pruebas

El entorno de pruebas es el que contiene los elementos propios de la prueba, el entorno tradicional utilizado para realizar las pruebas de forma manual.

El objetivo de este entorno de pruebas consiste en analizar la capacidad de un sensor de red de capturar y analizar tráfico en una red de alta velocidad. Para ello, cuenta con tres elementos principales, una máquina que se encarga de la inyección de tráfico parametrizado, otra máquina que actúa como receptor del tráfico, y una tercera máquina que se encarga de capturar y analizar el tráfico circulante por la red.

Hay un cuarto elemento, el switch, encargado de interconectar los tres elementos anteriores.

A continuación se describen más detalladamente cada uno de estos elementos.

- El inyector consiste en una máquina encargada de inyectar tráfico parametrizado. Algunos de estos parámetros incluyen la tasa de inyección, la longitud de los paquetes, el contenido de los datos...

- El sensor es una máquina multiprocesador encargada de capturar y analizar todo el tráfico que circula por la red, de forma concurrente. El objetivo de las pruebas consiste en observar y determinar el comportamiento de este sensor en diferentes configuraciones y con diferentes tipos de tráfico circulando por la red.

- El receptor consiste en una máquina encargada de recibir el tráfico introducido en la red por el inyector. No tiene mayor objeto que ser la máquina de destino del tráfico.

- El switch es el encargado de comunicar los tres elementos anteriores. Se trata de un switch Gigabit, que debe tener conectado el sensor en un puerto monitor, de tal forma que todo el tráfico que circula por la red sea visto por el sensor, independientemente de que sea o no su destinatario.

4.2. Entorno de gestión

El entorno de gestión se compone de los elementos interconectados a través del segundo switch, como se muestra en la Fig. 4. Así, además de los elementos del entorno de pruebas, cuenta con dos elementos nuevos, la máquina encargada de las funciones de gestión y el segundo switch que se encarga de interconectar los elementos en esta red.

La utilización de este segundo switch permite aislar ambos entornos, de tal forma que la red de gestión de las pruebas no afecte a éstas.

En estos elementos físicos se distribuyen todos los componentes de la arquitectura, es decir, el gestor, los agentes, los demonios y los formateadores. A continuación se describe cada uno de los componentes desarrollados, junto con su localización física y sus características principales.

4.2.1. El gestor

El gestor se encuentra situado en la máquina que tiene el mismo nombre, y es la máquina que será utilizada por el adminitrador de las pruebas. En esta máquina el administrador realiza la configuración, realiza las pruebas ejecutando el gestor y obtiene la salida filtrada y representada mediante el uso de formateadores.

4.2.2. Los agentes

En la plataforma desarrollada aparecen tres agentes diferentes, en las tres máquinas que participan en las pruebas.

- En el inyector se encuentra el agente que va a gestionar los diferentes inyectores de tráfico.
- El agente que se encuentra situado en el sensor es el encargado de gestionar la mayoría de los demonios, ya que es el elemento más importante de la red. Los demonios que se encarga de gestionar son el de configuración de parámetros de captura, el de selección del driver de captura, los diferentes sensores de captura, y los de estadísticas del switch y del sensor, tanto de captura como de consumo computacional.
- En el receptor se encuentra el agente que controla al demonio encargado de recoger las estadísticas del receptor.

4.2.3. Los demonios

Los demonios son los elementos más específicos y los más numerosos. Los demonios utilizados para la implementación real de la arquitectura son los siguientes:

- Configuración de parámetros de captura. Este demonio se encuentra en el sensor y se encarga de configurar diferentes parámentros del sistema operativo de la máquina que afectan a la captura y análisis de tráfico, como son el tamaño de las colas y los búferes del socket de captura. El funcionamiento de este demonio es muy sencillo, ya que sólo debe modificar los ficheros que contienen estos parámetros. Así, en el arranque del demonio, éste modificará la configuración, y en la parada del demonio restaurará la configuración inicial. Se trata por lo tanto de un esclavo. Este demonio no genera resultados ni estadísticas.
- Selección del driver de captura. Como parte del desarrollo del sensor, se han realizado diversas modificaciones a diferentes niveles, incluyendo modificaciones en el driver del controlador de red o en la parte del código del núcleo del sistema operativo encargada de la captura y gestión del tráfico de red. El objetivo de este demonio es el de utilizar en cada momento la configuración de cambios determinada para la prueba. Así, su función principal será la de cargar diferentes módulos del kernel. Se trata también de un demonio esclavo, que en el arranque carga los módulos seleccionados y en la parada descarga los módulos cargados previamente. Éste demonio devolverá las estadísticas específicas que generan los módulos cargados.
- Sensores de captura. Estos se encuentran en la máquina denominada "sensor", si bien el sensor es el software específico de captura y análisis de tráfico desarrollado. Hay diferentes sensores desarrollados, con diferentes características. Uno de estos sensores cuenta únicamente con la parte de captura y está orientando a analizar el rendimiento exclusivamente en esta parte. Otro sensor consiste en un sistema completo de captura y análisis de tráfico, descrito en [3]. De éste existen múltiples variantes, con diferentes aproximaciones a la resolución de la problemática planteada, por lo que interesa poder ejecutar varios de ellos incluso en la misma batería de pruebas. Todos estos sensores son esclavos. En la fase de arranque del demonio, éste arranca el sensor correspondiente. En la fase de terminación del demonio, éste detiene el sensor, que a su vez almacenará las estadísticas asociadas a la captura de tráfico. El agente recogerá estas estadísticas a petición del gestor.
- Inyectores de tráfico. Estos demonios se encuentran en la máquina denominada "inyector". Su tarea consiste en la inyección de tráfico parametrizado en la red, para que sea analizado por el sensor. Se trata del elemento que define la duración de la prueba, ya que esta finalizará cuando acabe la inyección de tráfico. Por ello, se trata de un maestro. En la fase de arranque comenzará a inyectar el tráfico, y la fase de finalización consiste en la finalización de la inyección. En ese momento, además, el demonio almacena los resultados de la inyección, que el agente enviará al gestor. Para las pruebas realizadas se utilizaron dos tipos de inyectores, con diferentes parámetros de configuración. Uno de ellos, orientado a validar un modelo analítico incluye los siguientes parámetros: número de paquetes a inyectar, tasa de envío de paquetes, proporción de paquetes con requerimientos de análisis, longitud media de los paquetes, desviación de la media y tasa de bytes por milisegundos. Otro de los demonios de inyección, utilizado para la inyección de paquetes con patrones de tráfico determinados cuenta con los siguientes parámetros: número de paquetes, longitud de los paquetes, patrón a inyectar y la proporción de paquetes que contienen ese patrón.
- Estadísticas del switch. Este demonio se encarga de recoger las estadísticas sobre el tráfico que ha circulado por el switch en el periodo de pruebas. Se trata de un esclavo que pueden situarse en

cualquiera de los elementos que intervienen en la prueba. En la fase de inicialización este demonio coge los valores iniciales de las estadísticas. En la fase de finalización recoge los valores finales y almacena las estadísticas usando los valores tanto iniciales como finales. En la fase de recolección de pruebas, el agente envía estos resultados al gestor.

- Estadísticas de consumo del sensor. Este demonio se encarga de recoger estadísticas sobre el consumo de recursos hardware: procesador, memoria, accesos a caché... y permite complementar la información obtenida mediante las estadísticas de captura de tráfico proporcionadas por otros demonios. Consiste en un esclavo que en la fase de arranque inicializa las variables a almacenar y en la fase de finalización obtiene los resultados finales y calcula y almacena los resultados, para que posteriormente, en la fase de recogida de datos, el agente los envíe al gestor.

4.2.4. Los formateadores

Los formateadores se van a encargar de procesar toda la información generada por los demonios, seleccionando los datos pertinentes y obteniendo formatos válidos para la representación.

Durante la implantación real de la arquitectura se han utilizado dos formateadores, uno de propósito más general orientado a la representación de los datos y otro más específico orientado a la obtención datos para la realimentación de parámetros de un módelo analítico que representa la arquitectura del sensor.

Representación de resultados El formateador para la representación de resultados es un programa que realiza el procesado de la información en base a un fichero de configuración. En este fichero de configuración, basado en tecnología XML, se definen las estadísticas a representar. Estas estadísticas pueden ser bien directamente los valores de las estadísticas almacenadas por el gestor, o bien cualquier combinación de esas estadísticas con otras, o con valores literales. Esta combinación consiste en las operaciones aritméticas clásicas. Esto nos permite obtener, por ejemplo, la tasa de inyección de tráfico en función del número de paquetes, la longitud por paquete, y el tiempo total de inyección. Una vez seleccionadas las estadísticas a representar y la combinación entre ellas, el formateador obtiene los valores a representar y los convierte a formato Gnuplot[4] y MS Excel, para la obtención de gráficas.

Obtención de parámetros para el modelo Este formateador es más específico y simplemente recoge las estadísticas necesarias para calcular los parámetros deseados.

5. Conclusiones

Se ha desarrollado una arquitectura genérica y fácilmente extensible para la automatización de pruebas de laboratorio en entornos de red heterogéneos con el objetivo de validar las hipótesis planteadas en el curso de los trabajos de investigación. Este desarrollo ha permitido reducir drásticamente los tiempos requeridos para la realización de estas pruebas, tradicionalmente realizadas de forma manual, y evitar en un alto grado la necesidad de repetición de las mismas debido a fallos humanos.

Además, la generalidad de la arquitectura asegura la facilidad a la hora de ampliar el soporte de entornos de pruebas, con lo que resulta sencilla su implantación en redes complejas y heterogéneas.

En el desarrollo de la arquitectura también se han detectado puntos de mejora en los que se debe trabajar. Uno de los problemas más habituales en la realización de las pruebas es que en alguno de los dispositivos que conforman el entorno de pruebas y gestión no tenga los elementos funcionando correctamente, es decir, que los agentes no se encuentren a la escucha, que los demonios no se encuentren instalados, etc. Este problema puede no aparecer hasta el momento de finalizar las pruebas, si se produce en el momento de finalización de los demonios o en la recogida de estadísticas, por lo que puede implicar una gran pérdida de tiempo. Para solucionar este problema se puede desarrollar una unidad de testeo que verifique el correcto funcionamiento de todos los elementos antes de lanzar las pruebas.

Otro problema detectado que debe solventarse en el futuro es la situación en la que falla un test dentro de una batería larga de pruebas. En este caso, habría que aislar la configuración de esa prueba en particular y repetirla. Esto requiere un trabajo manual considerable, por lo que se podría desarrollar otro módulo que se encargara de almacenar la configuración de la prueba particular de forma que pueda repetirse de forma sencilla.

Referencias

[1] F. Yergeau, T. Bray, J. Paoli, C. M. Sperberg-McQueen, E. Maler. “Extensible Markup Language (XML) 1.0 (Third Edition). Febrero de 2004.

[2] XSL Transformations (XSLT): http://www.w3.org/TR/xslt

[3] A. Ferro, I. Delgado, A. Muñoz, F. Liberal. “Multiprocessor Architecture for Passive Analysis of Network Traffic Focusing on Complex QoS Strategies”. Proceedings on IEEE International Conference on Communication . ICC’2005. Seúl (Corea del Sur), 16-20 Mayo 2005.

[4] Gnuplot plotting program: http://www.gnuplot.info

Impacto de la configuración de parámetros de la capa RLC en las prestaciones del servicio de acceso a Internet sobre UMTS

E. González Parada, J. M. Cano García y A. Díaz Estrella.
Departamento de Tecnología Electrónica. ETSI Telecomunicación.Universidad de Málaga.
Teléfono:952132757 e-mail:eva@dte.uma.es

Abstract *The main goal of this paper is the analysis of the interaction between TCP and the UMTS link-level layer (RLC) to determine the optimal parameters configuration for both protocols. For this proposal, a web browsing simulation framework has been considered. This framework involves a web client accessing to remote HTTP servers through the UTRAN and Internet. The impact of the RLC configuration parameters on the behaviour of the multiple TCP connections simulated under this framework is analysed, showing their effect on web traffic performance.*

1. Introducción

Uno de los principales objetivos contemplados en el desarrollo de los sistemas de tercera generación es la prestación del servicio de acceso móvil a Internet. Sin embargo, las aplicaciones de Internet poseen unos requerimientos de calidad de servicio muy específicos, que junto con la naturaleza del enlace radio, dominada por elevadas tasas de errores, un ancho de banda limitado y altos retardos, han hecho que la tarea de diseño de los protocolos que intervienen en la transferencia de información y las estrategias de gestión de los recursos radio constituya un aspecto fundamental en el desarrollo de las especificaciones de los sistemas de tercera generación. En los sistemas UMTS, la transferencia fiable de datos en el interfaz radio, basada en mecanismos ARQ (*Automatic Repeat Request*), es realizada por los protocolos que se encuentran en la capa RLC (*Radio Link Control*), mientras que el control de acceso al medio y las estrategias de gestión de recursos, son llevadas a cabo por la capa MAC (*Medium Access Control*) y las capas de control superiores, respectivamente. Todas estas capas conjuntamente gestionan los diferentes parámetros de calidad de servicio en el interfaz radio de UMTS, que son requeridos por las aplicaciones o por los usuarios. No obstante, esta claro que el concepto de sistema UMTS no comprende únicamente el interfaz radio, y por lo tanto la calidad de servicio debe ser entendida extremo a extremo. En esta tónica, aplicaciones de Internet, como las anteriormente mencionadas, requieren una transferencia de datos fiable extremo a extremo. Hoy en día, en las redes fijas este tipo de transferencia es provista por el protocolo TCP (*Transmission Control Protocol*), por lo que es muy probable que siga siendo el protocolo que se utilice en los sistemas de acceso radio.

TCP se diseñó para trabajar en redes cable donde la probabilidad error de paquetes es muy baja, y por lo tanto, cualquier pérdida de paquetes se identifica como un estado de congestión de la red que TCP debe tratar, disminuyendo la tasa de transmisión de paquetes a través de los mecanismos de control de congestión que implementa. La disminución de prestaciones que esto produce en entornos radio, donde las pérdidas son debidas a errores de transmisión, ha generado un enorme esfuerzo de investigación en los últimos años destinado al estudio de este problema y a la propuesta de multitud de soluciones diferentes [1, 2, 3, 4].

No obstante, los errores en el enlace radio no son la única causa que puede afectar a las prestaciones de TCP, sino que como se ha probado en estudios recientes [5, 6, 3, 7, 8, 9], los mecanismos de control que incluyen las capas más bajas de las redes radio también pueden interaccionar con TCP, degradando sus prestaciones. La incapacidad de TCP para distinguir entre las pérdidas causadas por congestión y las debidas al canal radio ha motivado, entre otras soluciones, el uso de mecanismos de retransmisión local que resuelven el problema del impacto de las pérdidas radio en las prestaciones de TCP, lo que en el caso de UMTS es proporcionado por la capa RLC. Sin embargo estos mecanismos de retransmisión introducen variabilidad tanto en el ancho de banda como en la latencia de la red debido al número también variable de retransmisiones a las que se ven sometidos los paquetes en el enlace radio. Estas fluctuaciones pueden causar una degradación en las prestaciones de TCP debido a la dificultad para seguir las variaciones a corto plazo que tienen los mecanismos que incorpora este protocolo para estimar la tasa de transmisión de datos de la fuente y el temporizador de retransmisión más adecuados al estado de la red. Por todo lo comentado anteriormente, el estudio de las prestaciones de TCP no se puede ver desvinculado de los mecanismos que subyacen y rigen la red, y por lo tanto a la hora de diseñar una red de acceso radio, éste diseño no se puede focalizar en el la capa de transporte o en el de las capas inferiores, sino que hay que tener en cuenta la interacción de los mecanismos que operan en

cada capa [9, 7].

2. Objetivos

Recientemente, la comunidad investigadora ha centrado su atención en analizar la interacción entre TCP y los protocolos que forman parte de la red de acceso radio UMTS (UTRAN), surgiendo diferentes alternativas a la hora de tratar el impacto que las pérdidas radio y la variabilidad en el ancho de banda y retardo de los paquetes tienen las prestaciones de TCP. Las diferentes aproximaciones optan habitualmente por actuar sobre una capa concreta proponiendo modificaciones y mejoras, aunque por lo general se deducen en base al análisis de la interacción entre capas. Así, existen estudios que proponen la modificación de los mecanismos y parámetros de TCP para mejorar su comportamiento bajo estas condiciones [10, 11, 12, 7]; mientras que otros estudios se centran en modificar los mecanismos y la configuración de los parámetros de las capas que integran la red de acceso radio, con el objeto de disminuir el impacto negativo que causan en las prestaciones TCP. Dentro del interfaz radio, existen también diferentes niveles de actuación, encontrándose en la literatura estudios que se centran en la capa MAC y los mecanismos de gestión de recursos [13, 14, 15, 16, 17, 18, 19, 20, 5, 5, 21, 22], y estudios que se ocupan preferentemente de la configuración de la capa RLC, estudiando el efecto de diferentes parámetros en conexiones TCP [23, 24, 25, 26, 27].

Dentro de esta última línea es donde puede englobarse el presente trabajo, que pretende completar a los estudios anteriormente citados, que habitualmente se limitan a medir el impacto de la configuración de la RLC en el retardo de paquetes y en el throughput para conexiones TCP de larga duración. En el presente artículo se propone ampliar el análisis mediante el estudio del impacto de los diferentes parámetros de la capa RLC en las prestaciones percibidas por los clientes web, que generan un tráfico con unas características más complejas. Para ello, no sólo será necesario analizar el efecto que los diferentes parámetros de configuración de la capa RLC de la red de acceso UTRA tienen sobre el rendimiento de los servicios ofrecidos sobre TCP, sino que además será necesario establecer la relación que dicho efecto tiene con la propia configuración de TCP y con las características del tráfico.

3. Investigación relacionada

La capa RLC en UMTS [28] establece los mecanismos necesarios para la fragmentación y reensamblado de los paquetes de datos de nivel superior (SDUs) en bloques de menor tamaño (RLC PDUs), de forma que puedan transportarse sobre el interfaz radio. La RLC puede configurarse en modo transparente, modo no confirmado o modo confirmado. En este último modo, que es en el que el presente artículo centra su atención, la capa RLC implementa un mecanismo ARQ para paliar los errores que se producen en el enlace radio y que se traducen en la pérdida por corrupción de bloques PDUs. Para ello, la RLC implementa un mecanismo de ventana deslizante con retransmisiones selectivas. De acuerdo con este mecanismo la entidad RLC receptora envía, bajo ciertas condiciones, informes de estado hacia la entidad transmisora, indicando qué PDUs se han recibido correctamente y cuáles deben ser reenviadas. En base a esta información que recibe de la entidad receptora, la entidad transmisora decide retransmitir determinadas PDUs o transmitir PDU nuevas si no hay nada que retransmitir. Para evitar bloqueos, la entidad transmisora puede solicitar a la receptora el envío de información de estado mediante la activación de un bit de *Poll* en los paquetes enviados. Al recibir una indicación de *poll*, la entidad receptora debe enviar inmediatamente un paquete con información de estado. La norma UMTS prevé diversos mecanismos y criterios para disparar la solicitud de *poll* en el transmisor, o el envío de información de estado por decisión del receptor, que dan lugar un cierto número de parámetros configurables. La descripción detallada de estos mecanismos puede encontrarse en [28] y cae fuera de los objetivos del presente artículo. No obstante, en la tabla 1 se muestra un resumen de dichos parámetros.

Además de los parámetros asociados a los mecanismos de solicitud y envío de mensajes de estado, existen otros parámetros relacionados con el descarte de información si no se consigue transmitir de forma correcta, que también se muestran en la tabla 1. Después de realizar un determinado número de retransmisiones sin éxito o transcurrido un cierto tiempo sin que una SDU se transmita adecuadamente, el transmisor RLC puede decidir descartar la SDU y comenzar a intentar transmitir la siguiente. Para ello, enviará una solicitud de desplazamiento de la ventana (MRW) al receptor, que este deberá confirmar para completar el descarte de la información. Esta solicitud se reenvía al cabo de un tiempo si el receptor no la confirma, para evitar bloqueos.

En general, no existen muchos estudios que se centren en la influencia de los mecanismos y configuración de parámetros de la capa RLC sobre el tráfico TCP, y habitualmente el análisis se limita a uno ó dos parámetros como mucho, realizándose siempre en un entorno donde se simulan una o muy pocas conexiones TCP. En concreto, en [23], se estudia la influencia de distintos parámetros de la RLC y del enlace sobre una única conexión TCP configurada de distintas formas. Estos parámetros son: el número máximo de retransmisiones permitidas en la capa RLC, el retardo del tramo radio,

Parámetros Transmisor	
Timer Poll Periodic	El envío de peticiones de *poll* puede dispararse en base a este temporizador
Poll PDU	El transmisor puede realizar una petición de poll cada vez que envía un número de PDU
Poll Window	El transmisor puede realizar una petición de poll cuando el número de PDUs no confirmadas supere un cierto porcentaje de la ventana máxima de transmisión
Timer Poll	Es un temporizador que dispara el reenvío de una petición de *poll* si no se ha recibido información de estado transcurrido un tiempo después de haber mandado una petición de *poll*
Timer Poll Prohibit	Es un tiempo de inhibición de envío de peticiones de *poll* después de haber enviado una
Timer SDU Discard	Es un temporizador para el descarte de SDUs si no se han conseguido transmitir con éxito transcurrido un cierto tiempo
SDU Discard	Las SDUs pueden descartarse después de haber retransmitido sin éxito las PDUs que la forman un cierto número de veces
Timer MRW Retransmit	Timer para el reenvío de una solicitud de descarte de SDUs transcurrido un tiempo sin recibir confirmación de un descarte previo
Parámetros Receptor	
Timer Status Periodic	Es un temporizador que puede disparar el envío de información de estado
Timer Status Prohibit	Es un tiempo de inhibición de envío de información de estado tras haber realizado un envío

Tabla 1: Parámetros configurables de la capa RLC

el patrón de error y el tamaño de la ventana TCP. Las conclusiones que se obtienen del estudio sostienen que el modo con reconocimiento de la RLC es el más aconsejable para las aplicaciones que usan TCP, que el número de retransmisiones de la RLC se debe establecer al máximo posible, que debido a la mejor reacción del protocolo TCP ante un patrón de pérdidas a ráfagas es ventajoso utilizar bajas profundidades de *interleaving*, y por último, en relación con el tamaño del buffer que es aconsejable maximizar el tamaño de la ventana TCP pero sin sobrepasar el producto ancho de banda retardo de la red. De igual forma, [24] analiza una única conexión TCP, estudiando los factores que afectan a la ocupación del buffer RLC. Según dicho estudio la ocupación del buffer RLC disminuye a medida que el retardo es mayor en el tramo de Internet y aumenta a medida que la tasa de error es mayor. Además se concluye que a medida que el buffer RLC aumenta también lo hace el *goodput* de TCP. En [25] se estudia un conjunto de conexiones TCP, pero haciendo uso de un modelo analítico de TCP donde no se modelan todos los mecanismos. Particularmente, se estudia el impacto en el *trhoughput* y retardo de la RLC de dos parámetros concretos, *Timer Poll Prohibit* y *Poll Timer*. A través de diferentes experimentos los autores proponen unos valores para ambos temporizadores, dejando abierto el estudio de la interacción con el resto de parámetros que constituyen la RLC. En [27, 29], los autores desarrollando modelos analíticos tanto del esquema ARQ utilizado en la capa de enlace, como de TCP, estudian la interacción entre TCP y el protocolo ARQ en función de las condiciones variables del canal y en presencia de diferentes requisitos de calidad de servicio. Como resultado proponen un algoritmo adaptativo a las características variables del enlace radio que permite establecer algunos parámetros de los protocolos ARQ y TCP. Los parámetros que son objeto de estudio son el número de retransmisiones del protocolo ARQ y el tamaño del segmento TCP. En un sentido similar, en [26], pero haciendo uso de de la implementación de un modelo de simulación de la RLC, los autores proponen un algoritmo para establecer de forma adaptativa el número máximo de retransmisiones de la capa RLC, sosteniendo que como, aunque sea pequeña, existe una probabilidad de error en el tramo cable, no es necesario hacer el enlace radio totalmente fiable. Tras la propuesta de este algoritmo y las diversas pruebas llevadas a cabo, los autores concluyen que si bien establecer un número limitado de retransmisiones mejora las prestaciones de TCP, está mejora no es significativa. Por ello, indican que lo más simple es adoptar un protocolo de la capa de enlace fiable, y por lo tanto con un número ilimitado de retransmisiones. El resto de los parámetros no son objeto de estudio en dicho trabajo.

En este contexto de investigación es donde tiene cabida el estudio realizado en el presente artículo, en el que se evalúan distintos valores y configuraciones de los parámetros de la capa RLC en un marco de evaluación más cercano a la realidad.

3.1. Escenario de evaluación

El escenario de evaluación se ha escogido de forma que sea lo más representativo y realista posible, sin aumentar en exceso el coste computacional de la simulación. Es por ello que, para analizar la influencia de la configuración RLC en las prestaciones del acceso móvil a Internet, se ha optado por simular un escenario consistente en una sesión de navegación web sobre una conexión de acceso establecida a través de un canal dedicado. En este sentido, es necesario indicar que el impacto de la configuración RLC podría ser distinto sobre otros servicios con características diferentes. No obstante, se ha descartado simular otros tipos de tráfico como tráfico *peer to peer* o conexiones de larga duración por su escasa adecuación a un entorno de acceso móvil a Internet en el que la tarificación no es plana.

Una vez seleccionado el escenario es necesario estudiar qué capas del sistema se pueden obviar en la simulación de forma que se reduzca el coste computacional de la misma sin afectar al objeti-

vo que se desea alcanzar. El modelo de simulación utilizado en el presente artículo se muestra en la figura 1. En este caso, es obvio que la interacción de la capa RLC con el protocolo de transporte TCP para un determinado tipo de tráfico obliga a la simulación detallada de estos tres elementos. En cambio, el comportamiento de otras capas como la capa de red IP y la de adaptación PDCP, así como los mecanismos de transporte por la red núcleo UMTS puede ser simplificado en un primer análisis. Finalmente, las capas MAC y física quedan caracterizadas por el número de bloques de transporte por TTI, el tamaño del bloque de transporte y la probabilidad de error de bit.

El esquema de transferencia de información es el que aparece recogido en la figura 2. En este esquema se muestra el transporte y tratamiento de información tanto para el enlace radio como para el tramo cable formado por la red núcleo e Internet. En particular se detalla el tratamiento de la información en la capa RLC, mostrando como la entidad transmisora de la RLC perteneciente al equipo servidor puede enviar PDU de datos correspondientes a los segmentos TCP, y PDU de estado correspondientes a la recepción de los ACK de dichos segmentos TCP. A su vez, la entidad receptora de la RLC del equipo servidor puede recibir PDU de datos que encapsulan los ACK de los segmentos TCP enviados, y PDU de estado relacionadas con las PDU de datos que envió la entidad transmisora RLC de este mismo equipo.

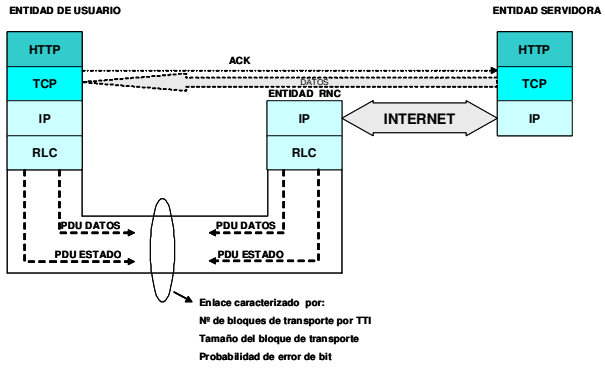


Figura 1: Escenario para la evaluación de la capa RLC

En lo que respecta a la generación de tráfico, en este artículo se ha utilizado un modelo estructuralista de tráfico web similar al propuesto en [30], donde el valor de los parámetros que definen el comportamiento del modelo han sido extraídos de las trazas reales capturadas en un entorno académico y corporativo [31]. En las simulaciones incluidas en el presente artículo se ha considerado una sesión web de duración indefinida en la que el cliente consulta páginas, cada una de las cuales consta de una serie de objetos que son descargados por el navegador mediante conexiones TCP. En las pruebas realizadas en este artículo se ha considerado que el navegador web mantiene un máximo de dos conexiones simultáneas para la descarga de objetos, lo que corresponde al comportamiento de los navegadores comerciales más ampliamente utilizados. Una vez descargada la página el usuario procederá a su lectura, por lo que se genera un tiempo de silencio antes de proceder a la descarga de la siguiente página. Los parámetros del modelo utilizado se muestran en la tabla 2. El análisis de la media y de los momentos de segundo orden revela que son muy similares a los obtenidos en caracterizaciones similares. Las variaciones detectadas en parámetros como el tamaño de la conexión se deben a que el modelo utilizado no trata separadamente los objetos que se encuentran en la cache, como sí se hace en [30].

Parámetro	Distribución	valores (media	desviación)
Tiempo entre páginas	Gamma	$\mu = 46,8s$	$\sigma = 168,6s$
Número de conexiones por página	Lognormal	$\mu = 5,3s$	$\sigma = 12s$
Numero de conexiones simultáneas	Fijo	2	
Tamaño de la conexión (descendente)	Pareto	μ = 5616 Bytes	$\alpha = 1,77$
Tamaño de la petición URI (ascendente)	Lognormal	μ = 364 Bytes	σ = 101Bytes

Tabla 2: Parámetros del modelo de tráfico de clientes web

El comportamiento de las conexiones TCP generadas por el modelo de tráfico para la descarga de los objetos es simulado detalladamente mediante un modelo bidireccional de TCP que incluye los principales mecanismos de dicho protocolo (establecimiento de conexión, slow start, congestion avoidance, fast retransmit, fast recovery, Reconocimientos selectivos, *Time Stamps*, etc.). La versión de control de congestión considerada ha sido la versión FACK [32], que utiliza la opción de reconocimientos selectivos y por tanto realiza la recuperación de pérdidas de forma más eficiente. Se ha utilizado esta versión de TCP porque cada vez son más los servidores que la soportan [33]. En la tabla 3 de la siguiente sección, se han incluido más detalles sobre la configuración de parámetros de TCP utilizada.

4. Pruebas y resultados

Antes de presentar las pruebas y los resultados obtenidos es necesario definir la métrica que se va a usar para la representación de los mismos. En la bibliografía relacionada con el presente artículo las métricas empleadas para la evaluación están basadas en el *goodput* y en el *throughput*. En esta ocasión, por el modelo de tráfico empleado, es interesante que la métrica de evaluación seleccionada dé información sobre la velocidad de transmisión efectiva de los datos a nivel de aplicación, y por lo tanto una idea de la percepción que tienen los usuarios de las prestaciones de las conexiones TCP en la red. Por este motivo, la métrica elegida va a estar basada en el *goodput*. En concreto, la

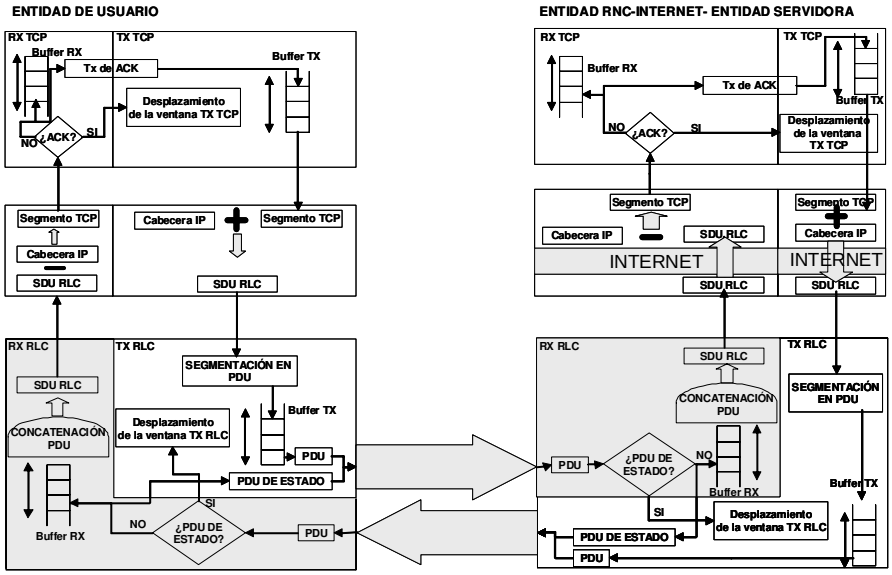


Figura 2: Esquema de transferencia de información en el modelo de simulación

media del *goodput* de página ponderado va a ser la medida establecida para la evaluación, ya que el usuario percibe la velocidad de transmisión de una conexión a través de la descarga de una página. Esta media, se obtiene a través de la ecuación 1, donde $S(p)$ y $T(p)$ son el tamaño y el tiempo de transmisión de la página p.

$$\widehat{A}_{Gpágina} = \frac{\sum_{p=1}^{P} S(p)}{\sum_{p=1}^{P} T(p)} \qquad (1)$$

Además de esta métrica, se han considerado también otras métricas como el promedio de los goodputs obtenidos por las diferentes páginas, la media ponderada del goodput de las conexiones y el promedio del goodput obtenido por cada conexión, pero los resultados no se han incluido por motivos de espacio y porque las conclusiones que pueden extraerse son las mismas que con la primera métrica propuesta.

Utilizando el escenario de evaluación descrito en apartado 3.1 cuyos parámetros de simulación por defecto son los que aparecen recogidos en la tabla 3, se han realizado una serie de pruebas encaminadas a determinar el impacto de la configuración de parámetros de la capa RLC en las prestaciones del servicio de acceso a Internet sobre UMTS. Para ello, se ha realizado un barrido de varios valores para diferentes parámetros de la RLC, dejando el valor por defecto para aquellos parámetros que no son barridos. Para cada una de las configuraciones analizadas se han ejecutado 5 simulaciones con diferentes semillas, en cada una de las cuales se ha simulado un total de 5 millones de conexiones TCP, generadas conforme al modelo de tráfico indicado en la sección anterior. El ancho de banda del acceso radio, que viene determinado por el TTI, el número de bloques por TTI y el tamaño de los bloques, es de 128 kbps.

Parámetros TCP	Valor
Tamaño del segmento	1460 Bytes
Tamaño máximo de la ventana	16 kBytes
Tamaño inicial de la ventana	2 MSS
ssthreshold inicial	16kBytes
Máximo retardo del ACK	0.2 s
Timeout Máximo	64 s
Timeout mínimo	400 ms
Timeout inicial	3 s
Granularidad del temporizador	200 ms
maxburst	4 MSS
Tipo	FACK
Parámetros la red de Acceso Radio	Valor
CAPA MAC-FISICA	
Tamaño del bloque de transporte	80 bytes
N° de bloques de transporte	4
TTI	20ms
CAPA RLC	
Tamaño PDU	78 bytes + 2 bytes cabecera
Tamaño de la ventana de transmisión	256 PDU
Tamaño del buffer de SDU	ilimitado
Timer Poll	2 TTI
Timer Poll Periodic	1.08s (54 TTI)
Timer Poll Prohibit	1TTI
Poll PDU	217 PDU
Poll Window	85 %
Timer Status Prohibit	Desactivado
Timer Status Periodic	Desactivado
Timer SDU Discard	Desactivado
Timer SDU Discard Retransmission	4 TTI
Parámetros de Internet	Valor
Retardo	50ms

Tabla 3: Parámetros de Simulación

Las figuras 3, 4, 5 y 6 muestran el efecto que los parámetros que determinan el envío de peticiones de *poll* por parte del transmisor tienen en las prestaciones observadas por los usuarios web, para diferentes tasas de error de bit en el medio radio. Los intervalos de confianza al 95 %, que no se representan en las figuras por claridad, son del orden de las decenas para todos los puntos. Las figuras 3, 4, 5 muestran el impacto de *Timer Poll Periodic*, *Poll Window* y *Poll PDU respectivamente*, mientras que 6 muestra la influencia de

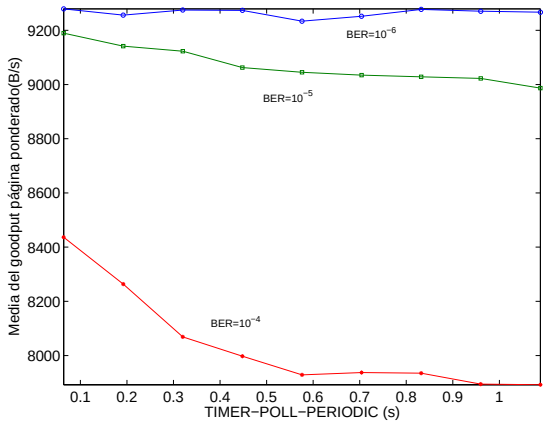


Figura 3: Efecto de la configuración de *Timer Poll Periodic* en las prestaciones del servicio web

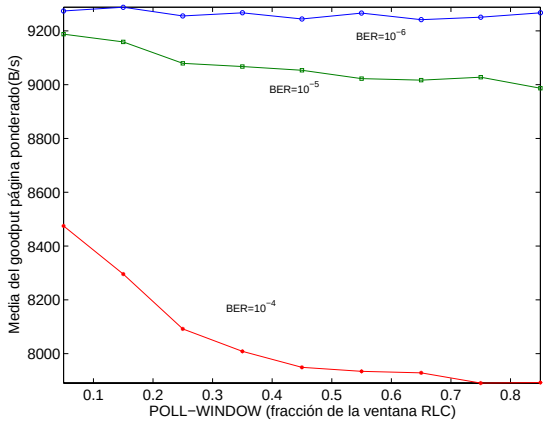


Figura 4: Efecto de la configuración de *Poll Window* en las prestaciones del servicio web

diferentes combinaciones de *Timer Poll* y *Timer Poll Prohibit* para $TimerPollPeriodic = 0{,}96s$, $PollPDU = 192$ y $PollWindow = 75\,\%$. En todas ellas se observa que no se producen grandes diferencias de prestaciones entre las distintas posibles configuraciones, aunque parece que las configuraciones de parámetros que provocan la petición de información de estado con una mayor frecuencia redundan en una mejora de las prestaciones. Como es lógico, este efecto es más acentuado conforme mayor es la tasa de error de bit (BER) en el enlace radio, ya que la pérdida de bloques por corrupción aumenta considerablemente.

Es conveniente señalar que todas las diferentes configuraciones de parámetros analizadas en el presente artículo obtienen unas prestaciones bastante aceptables. En este sentido, es necesario indicar que la configuración del *Timer Poll Periodic* no supera en ninguna prueba el 85 % del tiempo que tarda en transmtirse una ventana RLC completa, que *Poll PDU* nunca supera el 85 % de las PDUs de la ventana RLC y *Poll Window* no supera nunca el 85 %. Esto evita que se produzcan bloqueos y por tanto que se degraden las prestaciones de forma significativa. Si estos parámetros no se configuran de forma adecuada y consecuente con la ventana RLC establecida, sí que puede llegar a producirse una importante degradación de las prestaciones.

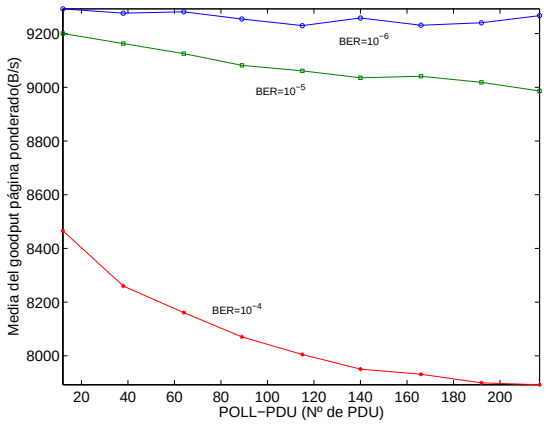
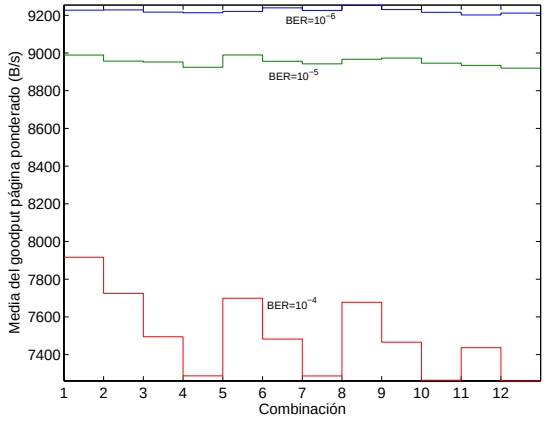


Figura 5: Efecto de la configuración de *Poll PDU* en las prestaciones del servicio web



	Poll Prohibit	Timer Poll		Poll Prohibit	Timer Poll
1	1 TTI	2 TTI	2	1 TTI	4 TTI
3	1 TTI	6 TTI	4	1 TTI	8 TTI
5	2 TTI	4 TTI	6	2 TTI	6 TTI
7	2 TTI	8 TTI	8	3 TTI	4 TTI
9	3 TTI	6 TTI	10	3 TTI	8 TTI
11	4 TTI	6 TTI	12	4 TTI	8 TTI

Figura 6: Efecto de la configuración de *Timer Poll* y *Timer Poll Prohibit* en las prestaciones del servicio web

Aunque su estudio se ha dejado fuera del presente artículo, cabe esperar la obtención de un resultado muy similar para los parámetros *Timer Status Periodic* y *Timer Status Prohibit* que son mecanismos de la entidad RLC receptora que pueden disparar el envío de información de estado, ya que su funcionamiento es similar a *Timer Poll Periodic* y *Timer Poll Prohibit*.

Por último, se ha procedido a estudiar la configuración del temporizador de descarte de SDUs, cuyo efecto para la distintas tasas de error de bit se muestra en la figura 7. Para estas simulaciones se ha establecido *Timer Poll Periodic=0.96s*, *Poll PDU=192* y *Poll Window=75 %*. El mecanis-

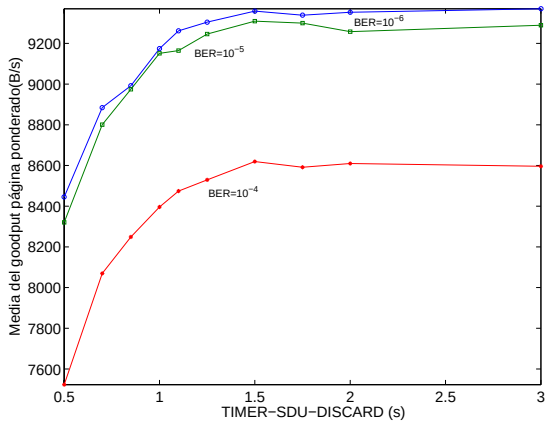


Figura 7: Efecto de la configuración de *Timer SDU Discard* en las prestaciones del servicio web

mo de descarte de la RLC promueve la eliminación de las SDUs que llevan mucho tiempo en el buffer sin haber podido ser transmitidas correctamente, al considerar que transcurrido un cierto tiempo dichas SDUs han podido ser retransmitidas a nivel de transporte (en este caso por TCP). Este mecanismo reduce el retardo que sufren las SDUs al atravesar el nivel radio, aunque evidentemente provoca la aparición de pérdidas de paquetes, que son enmascaradas por la capa RLC cuando este mecanismo no está activo. En el caso del tráfico que se estudia en el presente artículo, como se deduce de la figura 6, este mecanismo no introduce ningún beneficio, debido a que la aparición de pérdidas de paquetes en el enlace radio tiene un efecto bastante adverso en las prestaciones de TCP, lo que no se ve compensado en modo alguno por la reducción del retardo.

5. Conclusiones y líneas futuras

En este artículo se ha analizado el impacto de la configuración de la capa RLC de la red de acceso radio UMTS en las prestaciones del acceso a Internet a través de dicha tecnología. El estudio se ha centrado en las prestaciones de la navegación web, ya que es probable que este siga siendo el servicio de Internet más utilizado por los usuarios del acceso móvil a Internet. De este estudio se deduce que el funcionamiento de la RLC es aceptable para un conjunto bastante amplio de parámetros, si bien parece que para el tráfico analizado aquellas configuraciones que provocan el envío de informes de estado desde la entidad RLC transmisora a la entidad receptora con mayor frecuencia redundan en una mejora de las prestaciones del tráfico web. También se ha mostrado que la activación del mecanismo de descarte de SDUs para limitar el tiempo que los paquetes pueden permanecer en el buffer RLC reduce las prestaciones del servicio analizado.

El estudio que se ha realizado para el presente artículo puede ser extendido para cubrir el efecto de otros parámetros cuyo estudio no se ha incluido, y la posible interacción entre ellos. El estudio del efecto de la limitación del tamaño del buffer RLC y su interacción con el temporizador de descartes de SDU es una posible línea que será explorada en un futuro. Otro aspecto importante que debería estudiarse es la posibilidad de que el receptor envíe un informe de estado cuando detecte que falta un porcentaje de las PDUs que debería haber recibido. También es necesario completar el estudio para modelos de error más complejos, comprobando el comportamiento de la RLC con pérdidas a ráfagas.

Agradecimientos

Este trabajo ha sido parcialmente costeado por el proyecto de financiación pública Nº TEL2003-07953-C02-01.

Referencias

[1] H. Balakrishnan, V.Ñ. Padmanabhan, S. Seshan, y R. H. Katz, "A comparison of mechanisms for improving TCP performance over wireless links," *IEEE/ACM Transactions on Networking*, vol. 5, no. 6, pp. 756–769, 1997. [Online]. Available: citeseer.nj.nec.com/balakrishnan96comparison.html

[2] G. Fairhurst y L. Wood, "Advice to link designers on link automatic repeat request (ARQ)," IETF, RFC 3366, Agosto 2002.

[3] R. Chakravorty, S. Katti, J. Crowcroft, y I. Pratt, "Flow aggregation for enhanced tcp over wide-area wireless," in *IEEE INFOCOM*, Marzo 2003.

[4] K. Ratnam y I. Matta, "WTCP: An efficient transmission control protocol for networks with wireless links." *Proceedings of Third IEEE Symposium on Computers and Communications (ISCC '98)*, 1998.

[5] P. Ameigeiras, J. Wigard, y P. Mogensen", "Impact of TCP flow control on the radio resource management of WCDMA networks," in *VTC Spring 2002*, Mayo 2002.

[6] M. Methfessel, K. F. Dombrowski, P. Langendörfer, H. Frankenfeld, I. Babanskaja, I. Matthaei, y R. Kraemer, "Vertical optimization of data transmssion for mobile wireless terminals," *IEEE wireless Communications*, pp. 36–43, Diciembre 2002.

[7] M. Chan y R. Ramjee, "Improving TCP/IP performance over third generation wireless networks," in *IEEE INFOCOM*, Marzo 2004.

[8] P. Benko, G. Malicsko, y A. Veres, "A large-scale, pasive analysis of end to end TCP performance over GPRS," in *IEEE INFOCOM*, Marzo 2004.

[9] R. G. Mukhtar, S. V. Hanly, y L. Andrew, "Efficient Internet traffic delivery over wireless networks," *IEEE Communications Magazine*, pp. 46–53, Diciembre 2003.

[10] R. Chakravorty, J. Cartwright, y I. Pratt, "Practical experience with tcp over gprs," in *IEEE GLOBECOM*, Noviembre 2002.

[11] R. Ludwig y A. Gurtov, "Evaluating the eifel algorithm for TCP in a GPRS network," in *European WirelessNFOCOM*, febrero 2002.

[12] C. Casetti, M. Gerla, S. Mascolo, M. Sanadidi, y R. Wang, "TCP westwood: end to end bandwidth estimation for enhanced transport over wireless links," *Wireless Networks*, vol. 8, pp. 467–479, 2002.

[13] S. Heier, D. Heinrichs, y A. Kemper, "Performance evaluation of internet applications over the UMTS radio interface," in *VTC Spring 2002*, Mayo 2002.

[14] S. Heier, C. Ellerbrock, y M. Malkowski, "UMTS medium access control quality of service scheduling," in *IEEEPIMRC 2002 - Personal Indoor and Mobile Radio Conference*, vol. 5, Lisboa, Portugal, Septiembre 2002.

[15] S. Heier y M. Malkowski, "Umts radio resource management by transport format assignment and selection," in *Proceedings of WPMC 2002 - The 5th International Symposium on Wireless Personal Multimedia Communications*, Honolulu, USA, octubre 2002.

[16] M. Malkowski y S. Heier, "Interaction between UMTS MAC scheduling and TCP flow control mechanisms," in *ICCT - International Conference Communication Technology*, 2003, pp. 1373–1376.

[17] S. Heier, D. Heinrichs, y A. Kemper, "Ip based services at the umts radio interface," London, UK, mayo 2002.

[18] ——, "Performance of internet applications at the umts radio interface," in *Proceedings of 3Gwireless 2002 - 2002 International Conference on Third Generation Wireless and Beyond*, San Francisco, US, Mayo 2002.

[19] S. Heier, A. Kemper, y J. Rock, "Performance characteristics of umts for the mobile internet access," junio 2002.

[20] S. Heier, A. Kemper, S. Gräbner, y J. Rock, "Quality of service of internet applications over the UMTS radio interface," in *Mobile Multimedia Telecom Conference*, Junio 2002.

[21] A. Baiocchi y F. Vacirca, "End to end evaluation of WWW and file transfer performance for UMTS-TDD," in *IEEEGlobecom*, 2002, pp. 737–741.

[22] J. Chen y V. leung, "Improving end-to-end quality of service in 3G wireless networks by wireless early regulation of real-time flows," in *IEEE PIMRC 2003 - Personal Indoor and Mobile Radio Conference*, Beijing, China, Septiembre 2003.

[23] F. Lefevre y G. Vivier, "Optimizing UMTS link layer parameters for a TCP connection," in *VTC Spring 2001*, Mayo 2001, pp. 2318–2322.

[24] R. Bestak, P. Godlewski, y P. Martins, "RLC buffer occupancy when using a TCP connection over UMTS," in *PIMRC 2002 - Personal Indoor and Mobile Radio Conference*, vol. 5, Lisboa, Portugal, Septiembre 2002, pp. 2102–2106.

[25] X. Xu, Y. Chen, H. Xu, E. Gonen, y P. Liu, "Simulation analysis of rlc timers in UMTS systems," in *Winter Simulation Conference WSC 2002*, Mayo 2002.

[26] F. Vacirca, A. D. Vendictis, A. Todini, y A. Baiocchi, "On the effect of ARQ mechanisms on TCP performance in wireless enviroments," in *IEEEGlobecom*, Diciembre 2003.

[27] C. Chiasserini y M. Meo, "Impact of ARQ protocols on QoS in 3GPP systems," *IEEE Transactions on Vehicular Technology*, vol. 52, no. 1, pp. 205–215, Enero 2003.

[28] "RLC protocol specification," 3GPP, Technical Specification 25.322-v4.5.0 Release 4, Junio 2002.

[29] C. Chiasserini y M. Meo, "A reconfigurable protocol setting to improve TCP over wireless," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 6, pp. 1608–1620, Noviembre 2002.

[30] H. Choi y J. Limb, "A behavioral model of Web traffic," in *Proceedings of 7th IEEE International Conference on Networ Protocol (ICNP'99)*, september, 1999.

[31] A. Reyes-Lecuona, E. Casilari, A. Diaz-Estrella, A. García, y E. López, "Modelo estructural de tráfico WWW para la simulación de sistemas de acceso radio." *Ponencias de las VIII Jornadas de I+D en Telecomunicaciones*, 1998.

[32] M. Mathis y J. Mahdavi, "Forward acknowledgement: Refining TCP congestion control," in *Proceedings of the ACM SIGCOMM'96*, vol. 26, no. 4, 1996, pp. 281–291.

[33] A. Medina, M. Allman, y S. Floyd, "Measuring the evolution of transport protocols in the Internet," *ACM Computer communication Review*, vol. 35, no. 2, Abril 2005.

Ensamblado de ráfagas con diferenciación proporcional de servicios en redes OBS

Jairo Chapela Martínez, Manuel Fernández Veiga, Andrés Suárez González
Departamento de Ingeniería Telemática. Universidad de Vigo
ETSI de Telecomunicación. C/ Maxwell S/N. Campus Universitario.
36310 - Vigo (Pontevedra)
Teléfono: 986 81 21 00 Fax: 986 81 21 16
E-mail: jairo@det.uvigo.es

Abstract

In this paper, we address the issue of providing a proportional quality-of-service (QoS) for the existing classes of service in the traffic external to an OBS network. There are multiple references for achieving proportional QoS for the internal traffic of the OBS network —bursts—, but there don't exist any proposal for giving a proportional QoS scheme to the external traffic —packets that will be transported through the optical network—. This work offers a burst assembly algorithm for achieving the desired proportional QoS requeriments for external traffic. This algorithm is based on the existing methods for provide proportional QoS to internal traffic.

1. Introducción

La conmutación óptica de ráfagas (OBS, *Optical Burst Switching*) [1] promete ser una solución eficiente para el aprovechamiento del vasto ancho de banda que las fibras ópticas ofrecen para la transmisión de datos, eficiencia que se alcanza procesando electrónicamente sólo la información de control de los paquetes. Por el contrario, los datos viajarán del origen al destino en forma de señal óptica, sin sufrir ninguna conversión del dominio óptico al electrónico (ni viceversa) durante el trayecto.

Para compensar la lentitud relativa de los dispositivos electrónicos, los paquetes recibidos en la periferia de una red OBS se agrupan (ensamblan) [3] en unidades de mayor tamaño —denominadas ráfagas— que se conmutan por medios ópticos como una sola entidad después de que, en cada nodo, un paquete de control previamente transmitido haya preparado la configuración interna del conmutador óptico para el instante de llegada de la ráfaga. Esta reserva de los recursos de conmutación es necesaria por la inexistencia, hasta la fecha, del equivalente óptico de un búfer o memoria electrónica. Como el tiempo empleado en procesar la información de control se reparte entre todos paquetes de la ráfaga [2], la tasa efectiva de conmutación puede compararse a la de transmisión. Una red de OBS es, entonces, una red troncal de conmutadores sin búfer interno, a cuyos nodos están conectados conmutadores convencionales que se llaman de ingreso si ensamblan ráfagas e introducen tráfico en el dominio óptico, y de egreso si restituyen las ráfagas al dominio electrónico. Por tanto, en una red OBS es la contienda por los recursos de conmutación la causa primordial de la pérdida de tráfico, en lugar de la congestión.

La provisión de calidad de servicio (QoS) en una red OBS, tanto al tráfico interno —las ráfagas— como al tráfico externo —entre los nodos de ingreso y egreso— es un problema de claro interés que está recibiendo creciente atención en la literatura.

En particular, la provisión de QoS proporcional a las ráfagas de una red OBS ha recibido recientemente mucha atención [9][10][11]. En cambio, hasta donde nos es conocido, no hay trabajos que abordan la QoS proporcional para el tráfico externo. Este aspecto constituye, sin embargo, un modelo particularmente interesante para el tráfico TCP —el predominante en Internet— en vista de que, como concluyen los estudios presentados en [13] el *throughput* de una sesión TCP es proporcional a $p^{-1/2}$, en donde p representa la probabilidad de pérdida de un paquete. En consecuencia, relaciones proporcionales entre la probabilidad de pérdidas de los paquetes se traducen directamente en relaciones proporcionales entre el *throughput* de cada clase.

Este artículo aborda el problema de la diferenciación proporcional de tráfico externo en un número arbitrario de clases cuando la subred óptica posee algún mecanismo interno de diferenciación proporcional de la probabilidad de pérdida de las ráfagas. En concreto, el artículo propone y estudia un procedimiento para distribuir proporcionadamente los paquetes de las clases externas de tráfico entre dos o más clases de ráfagas, de forma que se satisfagan unas relaciones de proporcionalidad prescritas entre las probabilidades de pérdidas de los paquetes de las clases externas. Además, se deducen las condiciones analíticas que debe cumplir el mecanismo de diferenciación de las ráfagas para poder cumplir con las garantías de servicio proporcional que disfrutan los paquetes de cada clase.

Existe un trabajo relacionado [7], pero que no se plantea en ningún modo la proporcionalidad de la QoS.

La diferenciación proporcional por probabilidad de pérdida *en las ráfagas* es un modelo que ha sido estudiado en [8] donde, sin embargo, no se contempla cómo se ha de distribuir el tráfico de las clases externas para mantener la proporcionalidad.

El resto del artículo sigue esta organización. La sección 2 explica el funcionamiento del algoritmo de ensamblado de ráfagas. La sección 3 presenta el análisis matemático que conduce al cálculo del reparto de los paquetes entre las clases de ráfagas. La sección 4 presenta algunos resultados numéricos del rendimiento del mecanismo propuesto y la sección 5 recoge las conclusiones y extensiones posibles de este trabajo.

2. Descripción del algoritmo

2.1. Naturaleza del tráfico

Supóngase que a los nodos de ingreso de la red óptica llega tráfico previamente clasificado en diferentes niveles de prioridad, atendiendo a unos factores de proporcionalidad que relacionan las probabilidades de descarte de paquetes de unas clases con otras.

Por otra parte, en el interior de la red óptica se transmiten ráfagas de diferentes clases o prioridades, que por motivos de simplicidad, reduciremos a dos —ráfagas prioritarias y no prioritarias—. Las ráfagas serán tratadas de forma diferente en los nodos internos, dependiendo de la prioridad de las mismas, con objeto de diferenciar notablemente el servicio ofrecido a cada tipo de ráfaga. Así, las ráfagas pertenecientes a la clase prioritaria sufrirán una probabilidad de bloqueo en los conmutadores ópticos muy inferior a la sufrida por la clase menos prioritaria. Este grado de diferenciación será también proporcional y arbitrario.

El algoritmo aquí expuesto distribuirá los paquetes que entran en los nodos de ingreso en ráfagas de diferente prioridad durante la etapa de ensamblaje de las mismas. El objetivo buscado es el de garantizar el cumplimiento de las relaciones de proporcionalidad impuestas alterando la proporción de paquetes de una determinada clase de servicio que se integrarán en ráfagas de una determinada prioridad. Los aspectos relativos al funcionamiento del algoritmo se tratarán en detalle a continuación.

2.2. Ensamblaje de las ráfagas

Los paquetes entrantes en los nodos frontera pueden pertenecer a cualquiera de las N clases de tráfico —que, de ahora en adelante, denominaremos servicio intrínsecas—. En función de la clase a la que pertenezca un determinado paquete, éste se incorporará a una ráfaga prioritaria con cierta probabilidad, o de no darse el caso, en una ráfaga de prioridad baja.

La distribución de los paquetes en ráfagas puede llevarse a cabo siguiendo dos estrategias:

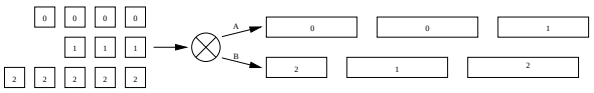


Figura 1: Distribución de paquetes en ráfagas homogéneas.

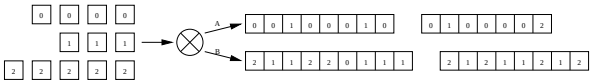


Figura 2: Distribución de paquetes en ráfagas heterogéneas.

- Componiendo ráfagas homogéneas (véase Fig. 1), es decir, haciendo que las ráfagas salientes de los nodos frontera sólo contengan paquetes pertenecientes a una sola clase intrínseca. Esta estrategia exigirá generar ráfagas prioritarias y no prioritarias en una proporción que garantice el cumplimiento de los factores de proporcionalidad impuestos a las clases de servicio intrínsecas.

- Ensamblando ráfagas heterogéneas (véase Fig. 2), pudiendo albergar una misma ráfaga paquetes de diferentes clases de servicio intrínsecas en la proporción adecuada para garantizar el cumplimiento de los factores de proporcionalidad exigidos.

2.3. Elección de la técnica de ensamblado

Habiendo presentado ya los dos modos de ensamblaje de ráfagas que se van a tratar en este estudio —homogéneas y heterogéneas—, procede ahora ver si, en base a todo el planteamiento anterior, son igualmente válidas las dos técnicas.

La forma en la que lleguen las ráfagas a los conmutadores ópticos de los nodos internos de la red tendrá repercusión en el funcionamiento del mecanismo de diferenciación de servicio. Por tanto, se atenderá a la distribución estadística que presentará el patrón de llegadas de las ráfagas a dichos nodos.

Un estudio realizado en [12] revela que, para un número suficientemente grande de longitudes de onda en cada enlace óptico, los procesos estocásticos de llegadas de ráfagas a los conmutadores de la red OBS se aproximan a una distribución exponencial. Dicho estudio asegura que en las citadas condiciones, es irrelevante la distribución del tamaño de las ráfagas, así como la del tiempo entre generación de ráfagas sucesivas. Dicho esto, se infiere que la elección de un criterio para la generación de las ráfagas —ya sea éste el vencimiento de un temporizador, el rebasamiento de un umbral para el tamaño de las mismas, o una combinación de estos dos—, es indiferente.

Bajo las condiciones del citado estudio, se asume que la llegada de ráfagas a los conmutadores se modela como un proceso de Poisson. Esto ocurre tanto para un algoritmo de ensamblado de ráfagas homogéneas como para uno de ráfagas heterogéneas. Si las condiciones del tráfico inyectado en la troncal óptica responde, de forma aproximada, a un mismo modelo estocástico sea cual sea el algoritmo utilizado para generar las ráfagas, es de esperar que se obtengan resultados similares para las dos técnicas contempladas.

3. Análisis matemático

Supóngase que el tráfico emergente a la entrada de los nodos frontera de la red óptica puede pertenecer a una de un total de N clases intrínsecas. Dichas clases se diferencian unas de otras en los factores de proporcionalidad en probabilidad de descarte de paquetes, s_i, que las caracteriza. Dichos factores cumplen que

$$s_1 = 1 < s_2 < \cdots < s_N.$$

Y la relación que liga los términos s_i con las respectivas probabilidades de descarte, Pd_i, viene dada por la expresión siguiente

$$\frac{s_i}{s_j} = \frac{Pd_i}{Pd_j} \quad \forall i, j = 1, \ldots, N. \tag{1}$$

Por otra parte, por la red óptica circularán ráfagas de tipo prioritario —que llamaremos de tipo A—, y ráfagas no prioritarias —de tipo B—. Los nodos internos de la red disponen de mecanismos para garantizar que la probabilidad de bloqueo para ráfagas de tipo A (P_A) es sensiblemente menor que la que corresponde a las ráfagas de tipo B (P_B).

$$P_A < P_B.$$

Si se conoce la proporción que relaciona P_A con P_B, es posible calcular el porcentaje de paquetes de cada una de las clases intrínsecas que se deberá incorporar en ráfagas prioritarias y no prioritarias.

3.1. Distribución de tráfico

Sean α_i y $\beta_i = (1 - \alpha_i)$ las probabilidades de que un determinado paquete perteneciente a la clase intrínseca i —habiendo un total de N clases intrínsecas—, sea ensamblado en una ráfaga prioritaria o no prioritaria, respectivamente.

La probabilidad de descarte de un paquete de clase i vendrá dada por la siguiente expresión

$$Pd_i = \alpha_i P_A + \beta_i P_B = (1 - \beta_i)P_B + \beta_i P_A.$$

Aplicando a este resultado la relación (1), se obtiene

$$s_i\big((1-\beta_j)P_A+\beta_j P_B\big) = s_j\big((1-\beta_i)P_A+\beta_i P_B\big). \tag{2}$$

Si se tienen N clases de servicio intrínsecas, se pueden definir $N - 1$ relaciones de proporcionalidad como las que siguen (cualquier otra relación de proporcionalidad posible entre una clase i y otra j es deducible de éstas)

$$\frac{s_2}{s_1}, \frac{s_3}{s_1}, \ldots, \frac{s_N}{s_1}.$$

Así, se plantea un sistema lineal de $N - 1$ ecuaciones con N incógnitas a partir de la expresión (2), donde cada una de las ecuaciones que componen el sistema son de la forma

$$s_i \Delta P \beta_1 - s_1 \Delta P \beta_i = (s_1 - s_i)P_A$$

con $\Delta P = P_B - P_A$.

El sistema de ecuaciones, expresado en forma matricial e introduciendo los cambios de notación $\mu_i = s_i \Delta P$ y $\phi_i = (s_1 - s_i)P_A$, queda como sigue

$$\begin{pmatrix} \mu_2 & \mu_1 & 0 & \ldots & 0 \\ \mu_3 & 0 & \mu_1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_N & 0 & 0 & \ldots & \mu_1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_N \end{pmatrix} = \begin{pmatrix} \phi_2 \\ \phi_3 \\ \vdots \\ \phi_N \end{pmatrix}.$$

Este sistema posee un grado de libertad, y por tanto, infinitas soluciones, que se expresarán en función de una de las incógnitas, como por ejemplo, β_1

$$\beta_i = \frac{s_i}{s_1}\beta_1 - \frac{(s_1 - s_i)P_A}{s_i \Delta P} \qquad \forall i = 2, \ldots, N. \tag{3}$$

Si ahora se considera que los nodos internos de la red OBS son capaces de ofrecer diferenciación de servicios proporcional en cuanto a la probabilidad de bloqueo sufrida por ráfagas de tipo A y de tipo B, se puede definir la siguiente relación de proporcionalidad

$$P_A = \gamma P_B \qquad ; \qquad 0 < \gamma < 1. \tag{4}$$

Combinando (3) y (4), se obtiene

$$\beta_i = \frac{s_i}{s_1}\beta_1 + \frac{(s_i - s_1)\gamma}{s_i(1 - \gamma)} \qquad \forall i = 2, \ldots, N. \tag{5}$$

Los factores β_i para las clases intrínsecas que van de la 2 a la N dependen del valor que se le asigne a β_1, que se corresponde con la clase 1 o más prioritaria. Para calcular este factor, se puede imponer la condición de que todos los paquetes pertenecientes a la clase intrínseca menos prioritaria crucen la red óptica en ráfagas de baja prioridad, esto es, $\beta_N = 1$. Así, a partir de la expresión (5) se deduce que

$$1 - \frac{s_N}{s_1}\beta_1 = \frac{(s_N - s_1)\gamma}{s_N(1 - \gamma)}. \tag{6}$$

De (6) se obtiene el valor de β_1 buscado

$$\beta_1 = \frac{(1 - 2\gamma)s_1 s_N - {s_1}^2\gamma}{{s_N}^2(1 - \gamma)}.$$

Los demás coeficientes β_i se calculan aplicando la fórmula (5).

3.2. Mecanismo de diferenciación interno

Considérese el caso extremo en el que todos los paquetes de la clase intrínseca menos prioritaria se destinan a ráfagas de baja prioridad ($\beta_N = 1$), y que los paquetes más prioritarios hacen uso solamente de ráfagas de prioridad alta ($\beta_1 = 0$). En esta situación, se deduce de la expresión (3) la siguiente igualdad

$$\beta_N = \frac{(s_N - s_1)\gamma}{s_N(1-\gamma)} \leq 1. \qquad (7)$$

A partir de (7) se puede inferir la condición que debe cumplir el mecanismo de diferenciación de un nodo interno para conservar la validez del planteamiento anterior

$$\gamma \leq \frac{s_N}{2s_N - s_1}.$$

4. Resultados de simulación

Una vez expuestos los conceptos y el formalismo matemático necesario para el cálculo de los parámetros que regularán el funcionamiento del algoritmo, se procede a demostrar la validez de éste mediante experimentos de simulación.

El escenario de prueba escogido es muy sencillo. Consta de un nodo interno al que se conectan tres nodos frontera mediante unos enlaces ópticos (ver Fig. 3). Cada uno de los nodos frontera inyecta en la troncal óptica ráfagas que contienen paquetes de tres clases de servicio intrínsecas diferentes. Los paquetes van dirigidos desde cada uno de los tres nodos frontera hacia los dos posibles destinos para ese origen concreto. En resumen, de cada uno de los tres nodos frontera salen tres flujos —uno de cada clase intrínseca—, hacia cada uno de los otros dos nodos frontera. En total, hay 18 flujos atravesando el nodo central de la red, contendiendo en grupos de 6 por cada posible salida de dicho nodo.

Los parámetros que dimensionan los enlaces ópticos son 64 longitudes de onda por fibra y 10 Mbit/s por cada canal.

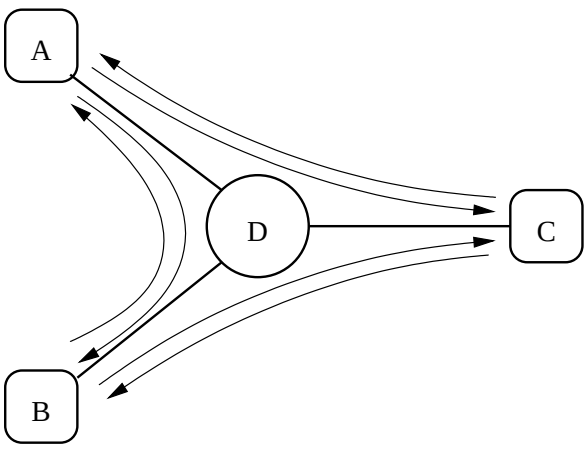


Figura 3: Escenario de prueba para la simulación.

4.1. Algoritmo de planificación

Aunque el planteamiento realizado en secciones anteriores es válido para cualquier mecanismo de diferenciación de servicios proporcional que se utilice en los nodos internos de la red OBS, es necesario seleccionar uno para realizar las simulaciones. Los resultados obtenidos mediante los experimentos de simulación se podrán extrapolar a casos en los que se utilicen mecanismos de diferenciación diferentes.

El algoritmo elegido para planificar cada una de las ráfagas salientes de un nodo interno de la red OBS combina una técnica de apropiación con un mecanismo de despriorización de ráfagas. Ambos tienen la misión de mantener el factor de proporcionalidad entre probabilidades de bloqueo de ráfagas de diferente tipo que se exige, manteniendo la estimación de éste invariable a las condiciones del tráfico ofrecido.

La apropiación permite descartar una ráfaga de baja prioridad ya planificada, para ceder los recursos que ésta tenía ocupados a una nueva ráfaga de alta prioridad. Esta técnica entrará en funcionamiento cuando sea necesario reducir la probabilidad de bloqueo de las ráfagas más prioritarias, hasta alcanzar la cota exigida.

Por otra parte, la despriorización de ráfagas consiste en planificar ráfagas prioritarias recibidas en los nodos internos como si se tratase de ráfagas de prioridad baja. Esto repercute en la probabilidad de bloqueo de las ráfagas de alta prioridad, acercándola a la probabilidad de bloqueo correspondiente a las ráfagas no prioritarias.

La combinación de las dos técnicas permite mantener la relación de proporcionalidad entre las probabilidades de bloqueo de ráfagas prioritarias y no prioritarias ante las variaciones en la carga ofrecida, de forma adaptativa. Mediante la técnica de apropiación se logra una diferenciación relativa de ambas probabilidades de bloqueo. Para ajustar dicho grado de diferenciación a las exigencias impuestas, se procederá a degradar —esto es, rebajar su prioridad— tantas ráfagas prioritarias como sea necesario, actuando de forma adaptativa, en función de las condiciones del tráfico ofrecido. A efectos de estimación del factor de proporcionalidad conseguido entre probabilidades de bloqueo, los nodos internos contabilizan los bloqueos sufridos por las diferentes clases de ráfagas, y en función de estos cómputos, alteran la conducta del algoritmo de planificación según lo dicho anteriormente.

4.2. Tráfico sintético

Para evaluar la respuesta del sistema a diferentes cantidades de tráfico ofrecido, se prueba primero generando tráfico de naturaleza sintética. El tráfico ofrecido se caracterizará por tener paquetes de longitud constante. Cada unidad tendrá un tamaño de 1250 bytes. Además, la llegada de los paquetes se comporta como un proceso de Poisson, con un valor

medio que se variará entre sucesivos experimentos, con el objetivo de estudiar el comportamiento del sistema a diferentes niveles de carga. Por último, el algoritmo de ensamblado de ráfagas opera por rebasamiento de un umbral de tamaño, disparando una nueva ráfaga cuando esta alcanza los 100000 bytes de longitud. La elección de estos valores no es arbitraria, ya que son magnitudes que reproducen los tamaños de paquetes y ráfagas más habituales. Dado, además, que tanto paquetes como ráfagas [12] respetan un patrón de llegadas con distribución exponencial, es sólo relevante, a efectos prácticos, el tamaño medio de paquetes y ráfagas, siendo indiferente su distribución estadística. Por este motivo, y en aras de facilitar las tareas de simulación, se asume una longitud constante para los paquetes.

En las simulaciones realizadas, la intensidad de tráfico ofrecido es diferente para cada clase intrínseca. La carga aportada por cada clase se ha escogido siguiendo un criterio que trata de emular una situación lo más realista posible. Según lo expuesto en [13], el caudal de salida máximo alcanzable por una conexión TCP guarda de forma aproximada una proporcionalidad inversa a la raíz cuadrada de la probabilidad de descarte de paquetes. Tratando de reproducir este hecho, la intensidad de tráfico ofrecido de cada clase intrínseca cumplirá dicha cualidad. Concretamente, se han simulado tres clases de tráfico intrínsecas, escogiendo como factores de proporcionalidad s_i los valores 1, 3 y 9. Por tanto, la intensidad de tráfico ofrecido para las dos clases más prioritarias es 3 y $\sqrt{3}$ veces la de la clase menos favorecida, respectivamente.

Con respecto a cada simulación llevada a cabo para la elaboración de este estudio, cabe decir que para cada situación se han hecho 10 experimentos de simulación, con no menos de 10^6 muestras en cada uno. Todo esto en virtud de conseguir que las estimaciones realizadas tengan un intervalo de confianza bastante pequeño, como es posible observar en las gráficas adjuntas.

4.2.1. Valores obtenidos para las pérdidas

La Fig. 4 muestra la probabilidad de bloqueo de ráfaga tanto para la clase prioritaria como para la no prioritaria para diversos valores de carga ofrecida por cada uno de los nodos frontera. Se puede observar que ambos valores crecen con la cantidad de tráfico inyectado en la red. Sin embargo, la relación entre ambas probabilidades de bloqueo se mantiene con independencia del tráfico ofrecido (dentro de un rango razonable).

Si ahora se observan los valores obtenidos para las probabilidades de descarte de paquetes pertenecientes a diferentes clases de prioridad intrínsecas, se aprecia nuevamente un crecimiento de estas con la carga de la red. No obstante, para un amplio rango de tráficos ofrecidos por cada uno de los nodos frontera, se mantienen las relaciones de proporcionalidad exigidas (véase Fig. 5).

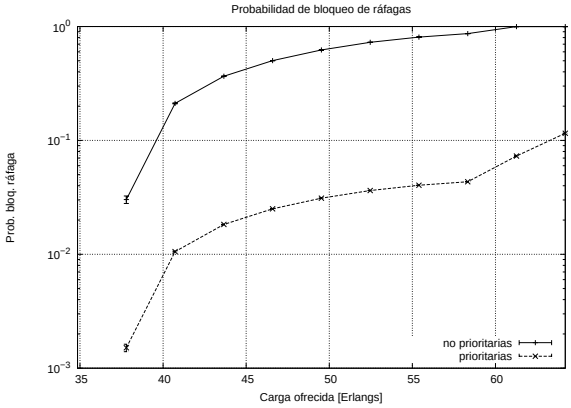


Figura 4: Probabilidad de bloqueo de ráfagas en un nodo interno.

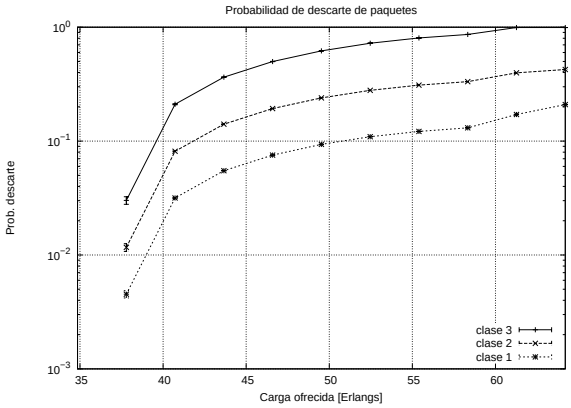


Figura 5: Probabilidad de descarte de paquetes.

4.2.2. Algoritmo de ensamblado

Como anteriormente se dijo, se contemplan dos variantes para la construcción de las ráfagas: homogéneas y heterogéneas. Según las hipótesis formuladas en el apartado 2.3.

Se han simulado ambos mecanismos para observar cual de ellos presenta mejor funcionamiento, y los resultados pueden verse en la Fig. 6. Como se puede observar, apenas se aprecia mejora alguna de un mecanismo frente a otro. Los factores de proporcionalidad se mantienen con ambos mecanismos de ensamblado, observándose un comportamiento un poco mejor a cargas bajas del algoritmo de ensamblado heterogéneo frente al homogéneo, en términos de probabilidades de descarte de paquetes.

El número de longitudes de onda necesarios para que ambos algoritmos ofrezcan prestaciones similares no es excesivamente alto. La Fig. 7 es el resultado de simular los dos algoritmos de ensamblaje estudiados, variando el número de longitudes de onda de los enlaces ópticos y, consecuentemente, la cantidad de tráfico ofrecido para mantener constante la carga. Dicha figura pone de manifiesto que, con pocas longitudes de onda, funciona mejor el ensamblado heterogéneo, aunque la diferencia no es muy significativa. A medida que se aumenta el número de longitudes de onda, el comportamiento observado al emplear cada uno de los dos algoritmos

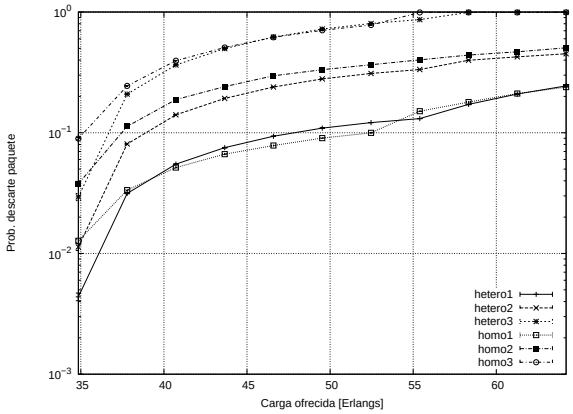


Figura 6: Comparación de los algoritmos de ensamblado (para 128 longitudes de onda).

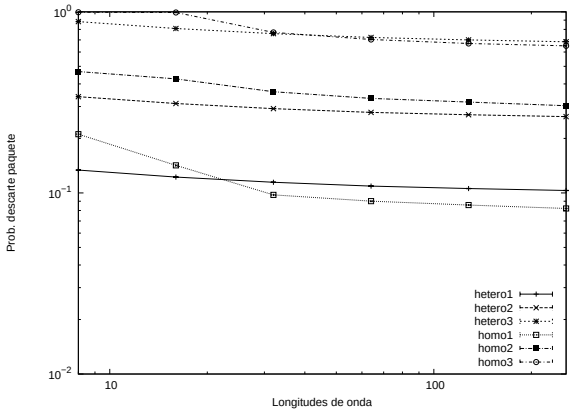


Figura 7: Influencia del número de longitudes de onda.

de ensamblado se hace muy similar.

4.2.3. Eficiencia global

Hasta ahora, no se ha aportado ningún criterio para la elección del valor óptimo de γ; tan sólo una cota superior para su valor. El siguiente experimento de simulación pretende buscar la influencia de dicho parámetro en la eficiencia del algoritmo, así como en la precisión con la que se consigue la diferenciación de las clases intrínsecas.

Mediante simulaciones, se ha estudiado la variación de la probabilidad de bloqueo de ráfagas en función de la carga ofrecida a la red por cada uno de los nodos frontera, para diversos valores de γ. La Fig. 8 ilustra dicha probabilidad de bloqueo, observándose un comportamiento más satisfactorio para los valores más grandes de γ, esto es, para aquellas situaciones en las que se exige una diferenciación menor a nivel de ráfaga.

5. Conclusiones

El ensamblado de ráfagas es un proceso propio de las redes OBS, y puede ser utilizado para ofrecer diferenciación proporcional de servicios a un conjunto de clases de tráfico externas a dicha red. Este

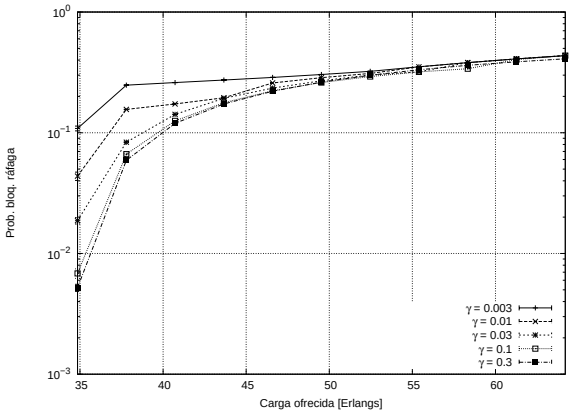


Figura 8: Probabilidad de bloqueo de ráfagas en función de γ.

artículo pretende dar una visión generalizada de los algoritmos de ensamblado de ráfagas, concluyendo que factores de los que gobiernan el comportamiento de dichos algoritmos son relevantes para el buen funcionamiento y la eficiencia de los mismos.

El análisis matemático sirve para calcular las proporciones de tráfico de cada clase intrínseca que deben integrar ráfagas prioritarias y no prioritarias, sean estas generadas de forma homogénea o heterogénea. También se revela la condición que debe cumplir el mecanismo de diferenciación residente en los nodos internos de la red, para garantizar el cumplimiento de las exigencias impuestas mediante los factores de proporcionalidad.

Los experimentos de simulación corroboran los resultados analíticos, para el caso concreto de utilizar un mecanismo de planificación de recursos con apropiación y despriorización de ráfagas, aportando además detalles acerca de la eficiencia del mismo.

Este estudio puede ser de utilidad para proveer diferentes niveles de calidad de servicio al tráfico emergente en la periferia de la troncal óptica. El mecanismo aquí descrito funciona para un nodo interno, siendo posible la extensión de su validez a un esquema de provisión de QoS extremo a extremo, introduciendo para ello los elementos arquitecturales que sean pertinentes. Como continuación de este trabajo, queda proponer los elementos que sirvan para adaptar a rutas completas los mecanismos que aquí son válidos para un único salto.

Agradecimientos

Este trabajo ha sido subvencionado por el Ministerio de Educación y Ciencia dentro del proyecto TIC2003-09042-C03-03 del Plan Nacional de I+D+I (parcialmente financiado con fondos FEDER) y por la Dirección Xeral da Xunta de Galicia con el proyecto PGIDT04PXIC32203PN (Plan Galego de Investigación, Desenvolvemento e Innovación Tecnolóxica).

Referencias

[1] T. Battestilli, H. Perros . "An Introduction to Optical Burst Switching". *IEEE Optical Communications*, vol. 41, núm. 8, págs. 10–15, ago. 2003.

[2] C. Qiao, M. Yoo. "Optical Burst Switching (OBS) - A New Paradigm for an Optical Internet". *Journal of High Speed Networks*, vol. 8, núm. 1, págs. 69–84, 1999.

[3] Y. Chen et al. "Optical Burst Switching: A New Area in Optical Networking Research". *IEEE Network*, vol. 18, núm. 3, págs. 16–23, may./jun. 2004.

[4] C. Dovrolis et al. "Proportional Differentiated Services, Part II: Loss Rate Differentiation and Packet Dropping". *Proc. Quality of Service, 2000. IWQOS. 2000 Eighth International Workshop on*, Pittsburgh, PA, págs 53–61.

[5] M. Yoo, C. Qiao, S. Dixit. "QoS performance of optical burst switching in IP-over-WDM networks". *IEEE J. on Sel. Areas in Comm.*, Vol. 18, núm. 10, págs. 2062–2071, oct. 2000.

[6] Y. Xiong et al. "Control Architecture in Optical Burst-Switched WDM Networks". *Journal on Selected Areas in Communications*, vol. 18, núm. 10, oct. 2000, págs. 1838-1851.

[7] V. M. Vokkarane, J. P. Jue. "Prioritized Burst Segmentation and Composite Burst-Assembly Techniques for QoS Support in Optical Burst-Switched Networks". *IEEE J. on Select. Areas in Comm.*, vol. 21, no. 7, págs. 1198–1209, sep. 2003.

[8] Y. Chen et al. "Providing Proportionally Differentiated Services over Optical Burst Switching Networks". *Proc. of IEEE Global Telecommunication Conference*, 2003.

[9] Y. Chen et al. "Proportional QoS over OBS networks". *Proc. of IEEE Global Telecommunication Conference*, 2001, vol.3, págs. 1510–1514, nov. 2001.

[10] H. C. Cankaya et al. "A preemptive scheduling technique for OBS networks with service differentiation". *Proc. of IEEE Global Telecommunication Conference*, 2003, vol. 5, págs 2745–2749, dic. 2003.

[11] Chee-Wei Tan et al. "Achieving proportional loss differentiation using probabilistic preemptive burst segmentation in optical burst switching WDM networks". *Proc. of IEEE Global Telecommunication Conference*, 2004, vol. 3, págs 1754–1758, dic. 2004.

[12] D. Morató et al. "Blocking time analysis of OBS routers with arbitrary burst size distribution". *Proc. of IEEE GLOBECOM 2003*, págs 2488–2492.

[13] J. Padhye et al. "Modeling TCP Reno Performance: A Simple Model and Its Empirical Validation". *IEEE/ACM Transactions on Networking*, vol. 8, No. 2, Abril 2000, págs 133–145.

Modelado y simulación en VHDL de una red OBS

Miguel Vicario y Sergio Lopez-Buedo
Grupo de Redes, Escuela Politécnica Superior
Universidad Autónoma de Madrid
C/ Francisco Tomás y Valiente, 11
28049 Madrid, España
e-mail: sergio.lopez-buedo@uam.es

***Abstract**. OBS is one of the most promising technologies for the next generation of Optical Internet. It supports IP over WDM, providing high percentage utilization of the wavelengths and, at the same time, offering extremely low buffering. In this paper an OBS network model has been built using the VHDL hardware description language. Its main purpose is to support future FPGA implementations of the control section of OBS nodes. It has also proved useful as a testbench to empirically analyze the impact of the variations of different parameters and topologies in the overall network performance. Studying these variations allowed us to make some improvements, like assigning a random wavelength and offset to the bursts in the edge nodes.*

1 Introducción

Actualmente el tráfico de voz no es tan predominante como hace unos años. Cada año el ancho de banda necesario para Internet se duplica [1], con lo que las nuevas tecnologías de transporte tienen que aprovechar cada vez más y mejor el ancho de banda, adaptándose a las peculiaridades del tráfico IP.

Con los recientes avances en tecnologías ópticas, especialmente en la multiplexación por división de longitud de onda (WDM), la cantidad de ancho de banda disponible en los enlaces de fibra óptica se ha incrementado notablemente. Mientras tanto, y de forma paralela, la ubicuidad del protocolo IP ha ocasionado que el paradigma generalista de IP sobre WDM sea considerado como la arquitectura óptima para la siguiente generación de Optical Internet [2].

Optical Burst Switching (OBS) es uno de los más prometedores mecanismos para soportar IP sobre WDM. Se sitúa en un punto intermedio entre la conmutación de circuitos y la de paquetes, combinando así las ventajas de la conmutación por multiplexación de longitud de onda con las del Optical Packet Switching (OPS).

La arquitectura de una red OBS está compuesta por dos tipos de nodos, los nodos frontera (EN, Edge Nodes) y los nodos troncales (CN, Core Nodes). Los primeros son los encargados de convertir el tráfico IP en ráfagas de datos OBS y viceversa, mientras que los segundos son los encargados de llevar a cabo las funciones de encaminamiento de estas ráfagas dentro de la red óptica.

El funcionamiento de una red OBS es conceptualmente sencillo. La reserva de los recursos es unidireccional (no hay confirmación), y se realiza mediante un paquete de control (BCP, Burst Control Packet*)* que viaja por una longitud de onda específica desde los nodos frontera hasta los nodos troncales. Cuando éstos reciben el paquete de control lo procesarán electrónicamente para posteriormente devolverlo al dominio óptico y reservar así los recursos en los siguientes saltos. Este BCP deberá ser enviado previamente a los datos con un offset predefinido, suficiente como para permitir el procesado opto-electro-óptico y realizar la reserva de recursos en todos los saltos de la red.

En la Fig. 1 se representa el esquema de bloques de un nodo troncal. El demultiplexor es el primer componente de la conmutación OBS: se encargará de separar el canal de control de entrada (ICC), usado por los BCPs, de los canales de datos (IDCs), utilizados por las ráfagas de datos. Cuando un BCP alcanza un CN, se convierte inmediatamente al dominio electrónico por el módulo de entrada (IM), y se redirige al router BCP, que determinará a qué fibra de salida se deben encaminar los datos.

Se pueden utilizar fibras con retardo de línea (FDLs), para retrasar la ráfaga de datos un tiempo igual al necesario para procesar electrónicamente la información del BCP, de tal manera que el offset entre ambos se mantenga constante. En cuanto el BCP haya sido procesado se transferirá al módulo de

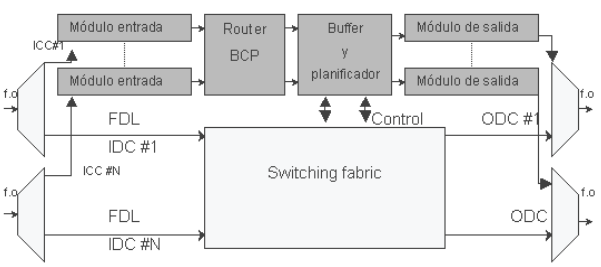


Figura 1: Arquitectura de un nodo troncal OBS

transmisión de salida (OM/TX), que actualizará los campos de control contenidos en el BCP y reservará la ráfaga de datos en una longitud de onda elegida por el planificador. Por último, el multiplexor insertará el BCP en la fibra de salida.

El trabajo aquí expuesto se ha centrado en realizar el modelado y simulación de una red OBS en el lenguaje de descripción hardware VHDL, con el objetivo de que permita investigar las posibles alternativas de implementación en FPGA de un nodo de control OBS. Se han modelado diversas topologías de red con diferentes valores de sus parámetros de funcionamiento.

Los resultados de estas simulaciones no sólo han sido útiles para demostrar la validez del modelo, sino que también prueban que gran parte de las pérdidas experimentadas en una red OBS son provocadas por las limitaciones de cada uno de los protocolos.

La primera de las arquitecturas que se ha modelado es el protocolo OBS originalmente propuesto por Qiao [3], en la que únicamente se puede producir conmutación en la longitud de onda, pero no se disponen de FDLs para retrasar los ráfagas y evitar así colisiones.

Por otra parte se ha considerado el protocolo Time Sliced Optical Burst Switching (TS-OBS) propuesto por J. Ramamirtham y J. Turner [4], en el que únicamente se pueden producir conmutaciones en el tiempo, no siendo posibles las conmutaciones en longitud de onda.

En ambos casos, las limitaciones llevan a descartar ráfagas en condiciones de baja carga, bien por no poder conmutar en el tiempo o en la longitud de onda. Como se muestra en experimentos posteriores, si se realiza una aleatorización de la posición de envío de los datos se obtiene una importante mejora en el rendimiento de estos protocolos.

2 Implementación

El primer modelo sobre el que se ha trabajado se muestra en la Fig. 2. Como se puede ver, la red está compuesta por 4 nodos troncales, cada uno de ellos con dos fibras de entrada y dos de salida. Cada fibra tiene W longitudes de onda para canales de datos y una para control, siendo W un parámetro modificable por el usuario.

Esta red es muy sencilla, pues el objetivo principal sólo era validar el modelo. La complejidad está en los propios nodos, de tal manera que el modelo es escalable de una manera muy simple a topologías más realistas. Esta implementación, como es lógico, sólo se ha centrado en la longitud de onda de control de cada nodo, que es la única que se convierte a eléctrico, y por tanto tiene interés en el diseño hardware.

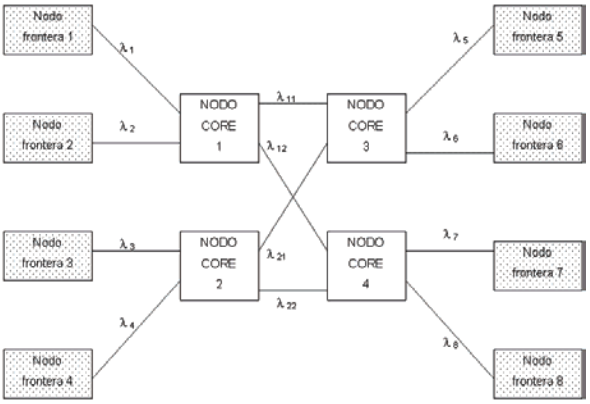


Figura 2: Topología de la red OBS modelada

Respecto a las herramientas usadas para la simulación, se ha empleado el simulador Active-HDL 5.2, de Aldec. Para la generación de números aleatorios se ha utilizado un paquete VHDL [5] que proporciona funciones con las distribuciones más frecuentes (poisson, normal, exponencial, uniforme, etc...).

2.1 Modelo de un nodo frontera

Los nodos frontera son los encargados de convertir el tráfico IP en ráfagas OBS y viceversa. En esta implementación son además los encargados de realizar la simulación IP.

Como es habitual, se ha supuesto que el tiempo de espera entre paquetes IP sigue una distribución exponencial y el tamaño de estos paquetes sigue una distribución de Poisson. Para el destino del tráfico se ha probado tres casos: con funciones de Poisson, Uniformes y dejando el destino constante, para ver el incremento en la tasa de descarte derivado de la saturación de los nodos troncales que enrutan hacia ese destino. El resto de parámetros de los nodos son completamente reconfigurables, y se han variado en los siguientes experimentos para obtener los valores óptimos de cada uno de ellos, y comprobar así el correcto funcionamiento del nodo en todos los casos.

Como ejemplo, se indican los siguientes valores:

- Tamaño de la tabla de reservas: 250 (en los nodos frontera 150)
- Tiempo entre slots: 40ns
- Número de longitudes de onda: 25
- Tamaño de la cola de entrada de peticiones: 20000
- Media de espera de los ráfagas: 20
- Media del tamaño de los ráfagas: 15

La tabla de reservas es la estructura donde se almacenan las ráfagas de datos, su longitud se mide en slots o unidades de tiempo elementales, iguales al tiempo de un BCP. Tiene tantas filas como longitudes de onda. En la Fig. 3 se pueden apreciar cada uno de

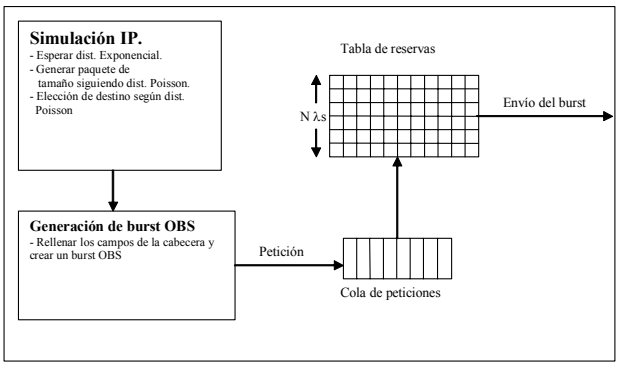


Figura 3: Bloques del modelo de un nodo frontera OBS

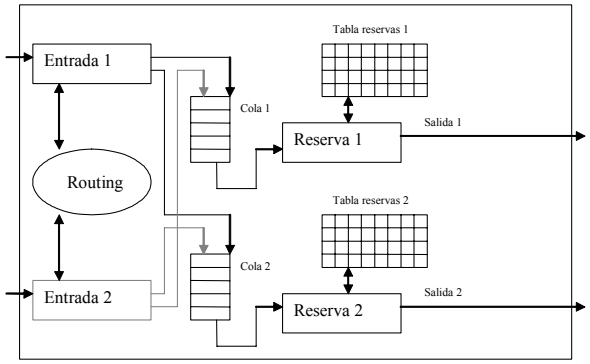


Figura 4: Bloques del modelo de un nodo troncal OBS

los diferentes módulos que componen un nodo frontera. Es este nodo el que realiza la simulación IP para generar tráfico siguiendo unos parámetros estadísticamente realistas. Una vez creado el tráfico IP se generará la ráfaga OBS asociada. En estos experimentos sólo se generan los parámetros estrictamente necesarios para el correcto funcionamiento del protocolo OBS, obviando campos de menor interés de las cabeceras del BCP.

Una vez generada la ráfaga de datos OBS se encolará para que otro proceso, el encargado de las reservas, extraiga progresivamente peticiones en la cola y marque determinadas posiciones en la tabla de reservas como ocupadas, en caso de efectuarse la reserva, o descarte la ráfaga en caso de que la tabla de reservas se encuentre llena.

Finalmente, se enviará el BCP por la longitud de onda de control, y con un cierto offset, la ráfaga de datos (aunque como ya se indicó antes, este modelo sólo tiene en cuenta la sección de control)

2.2 Modelo de un nodo troncal

Los nodos de core o nodos troncales son los encargados del encaminamiento de los bursts desde el nodo origen al nodo destino. Estos nodos son los que contienen las matrices de conmutación óptica junto con los cambiadores de longitud de onda y/o las fibras de retardo. El algoritmo de reservas utilizado es estático, y no tiene en cuenta ni las condiciones de carga del resto de nodos ni cualquier otra condición que influya en la tasa de descarte de las ráfagas.

Como ya se ha indicado, cada nodo troncal tiene dos entradas y dos salidas. Hay un proceso independiente por cada entrada, que es el encargado de recibir las peticiones por longitud de onda de control y de procesarlas. Cada uno de estos procesos leerá el destino del paquete y, mediante el proceso encargado del encaminamiento, decidirá por cuál de las dos colas debe mandar esa ráfaga, dependiendo del destino. Una vez metida la petición en la cola de reservas, un proceso (por cada tabla de reservas) extraerá la petición de la cola e intentará efectuar la reserva en la tabla adecuada. Si no es posible la ráfaga será descartada. Si, por el contrario, hay espacio disponible, se marcarán las posiciones de la tabla como ocupadas y se enviará la ráfaga por la longitud de onda correspondiente hacia el siguiente nodo.

2.3 Mecanismo de reserva

La técnica conocida como rellenado de huecos o void filling se basa en considerar los tiempos de inicio y de fin de los bursts aceptados en la reserva, consiguiendo así una mayor eficiencia dado que una nueva ráfaga puede ser reservada en un hueco libre, siempre que exista. Esta aproximación se engloba dentro de las llamadas RFD (Reserve a Fixed Duration), ya que el canal se reserva por una duración fija, correspondiente al tiempo de transmisión de la ráfaga, y son las que menores tasas de descarte obtienen.

Uno de los protocolos que siguen el método RFD es la del mecanismo de reserva Just Enough Time (JET). Este mecanismo fue propuesto por Qiao y Yoo [1] y añade, en la información de estado, los tiempos de comienzo y de fin de todas las ráfagas aceptadas, lo que implica una complejidad extra. Pero en contraste con Horizon [6], que sólo guarda el tiempo de finalización de las ráfagas, JET es capaz de detectar situaciones en las que no hay conflictos de transmisión aunque el tiempo de comienzo de la nueva ráfaga sea anterior que el de finalización de una previamente aceptada, ya que esta puede ser transmitida en un hueco entre dos ráfagas previamente reservadas.

Qiao y Yoo no sólo se quedan en lo anterior, sino que lo extienden para que sea capaz de soportar diferentes clases de QoS. Para ello, el offset de un burst de datos no sólo tiene un componente básico que representa la suma de los tiempos del proceso del paquete de control, sino que tiene un componente extra, llamado QoS offset, específico a cada clase de servicio. Como las ráfagas con offsets mayores experimentan menor bloqueo, los valores grandes de offset se asignan a las clases prioritarias

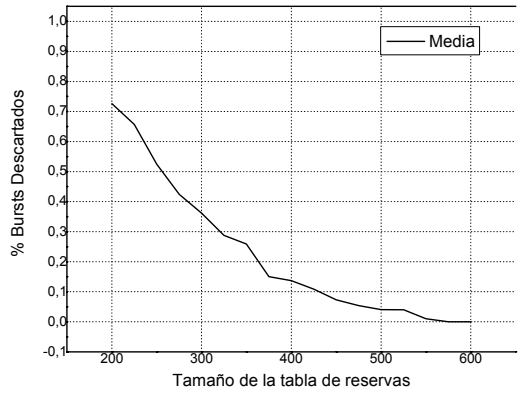


Figura 5: Variación del tamaño de la tabla en un nodo troncal

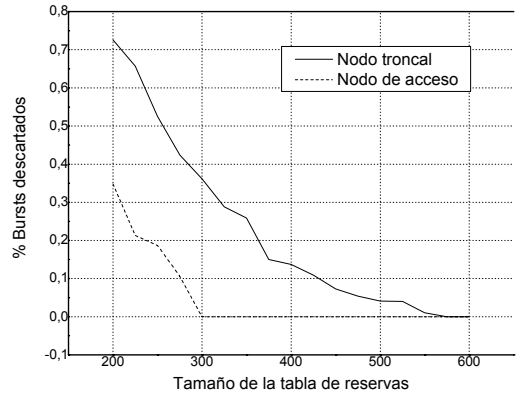


Figura 6: Comparación entre el rendimiento de un nodo frontera y uno troncal

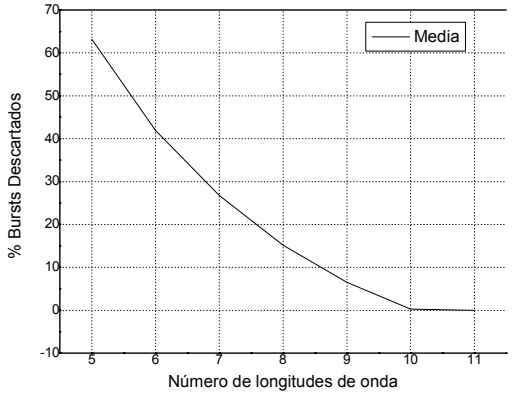


Figura 7: Variación del número de longitudes de onda en un nodo troncal

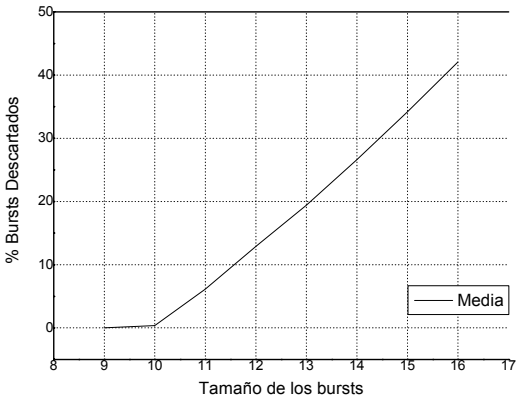


Figura 8: Variación del tamaño de los bursts en un nodo de core

3 Experimentos realizados

En primer lugar se comprobó la exactitud de la generación del tráfico IP, para después definir el valor de los parámetros de los nodos frontera de forma que no se descarte tráfico.

A continuación se comprobó el correcto funcionamiento de los nodos troncales (y consiguientemente de la red OBS en su totalidad) realizando una serie de experimentos con el objetivo de medir el rendimiento de la red ante la variación de diversos parámetros, como son el tiempo de simulación, el tamaño de los paquetes, el número de longitudes de onda disponibles, la distribución en la elección del destino de los paquetes y el tamaño de la tabla de reservas.

Finalmente, se presentan unos experimentos que muestran la importancia del algoritmo que se emplee al planificar la emisión de las ráfagas en los nodos frontera.

3.1 Variación del tamaño de la tabla de reservas

En este experimento se variaron conjuntamente el tamaño de la tabla de reservas de los nodos troncales. El objetivo de esta prueba es el de comprobar que el aumento del tamaño de la tabla de reservas es inversamente proporcional al número de descartes.

Para este experimento se fijó el tamaño medio de los paquetes en 10, al igual que el número de longitudes de onda disponibles y el tiempo de espera entre paquetes. En la Fig. 5 se observa, como era de esperar, que al aumentar el tamaño de la tabla de reservas disminuye la tasa de descarte.

Por otro lado, en la Fig. 6 se puede observar que a igualdad de tamaño en la tabla de reservas, los nodos frontera descartan menos ráfagas, por la simple razón de que los nodos troncales reciben el tráfico de dos nodos frontera.

3.2 Estabilidad de la red

Se ha comprobado que la red se estabiliza tras un periodo de tiempo determinado (variable según

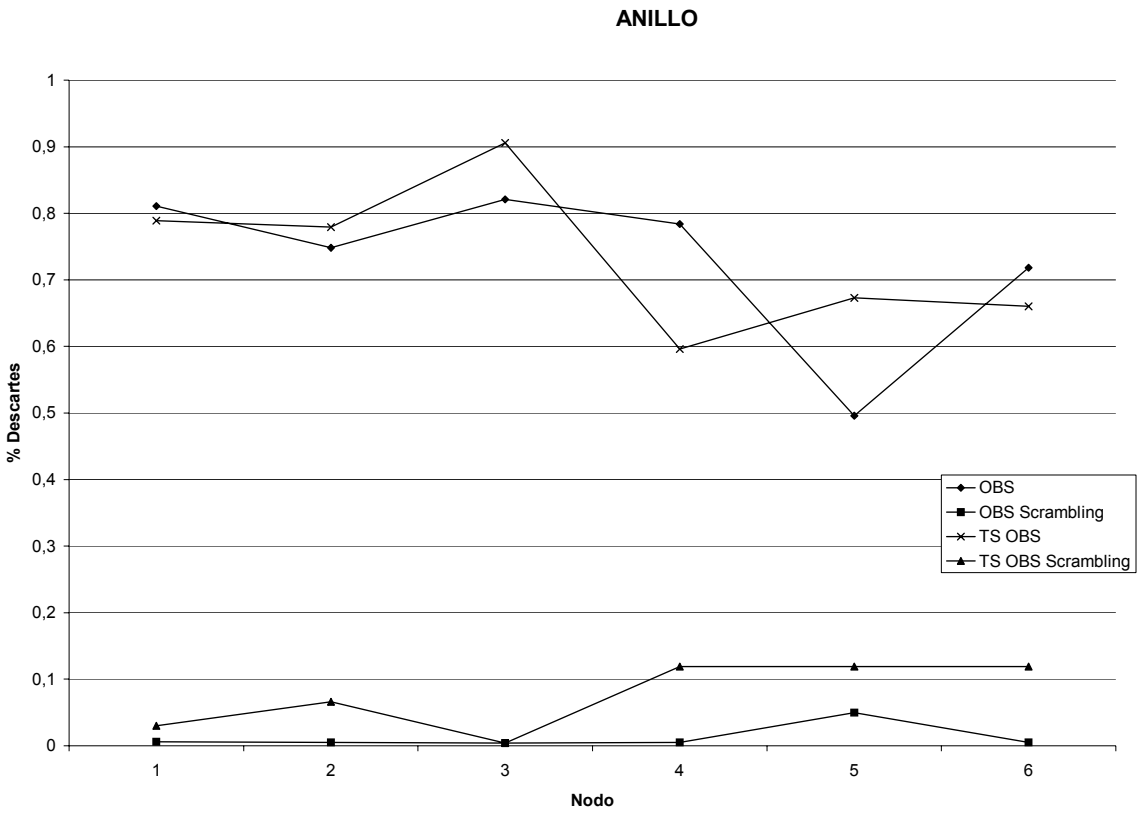


Figura 9: Tasa de descartes con y sin aleatorización de la posición de las ráfagas de datos en la tabla de reservas

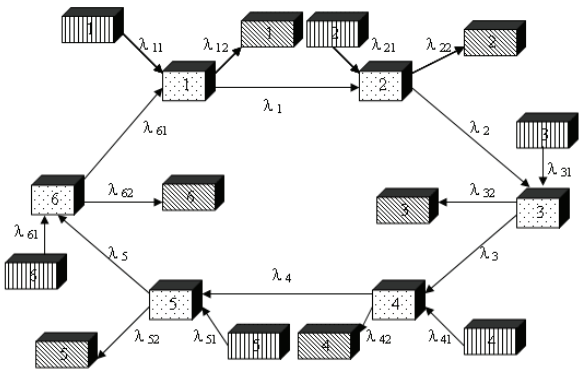


Figura 10: Topología de red en anillo

condiciones de carga), e incluso que existe un tamaño de las tablas de reservas (15 longitudes de onda y 400 posiciones temporales) para el que no existe ningún descarte de ráfagas, ni en nodos troncales ni en nodos de acceso, en un periodo de simulación muy prolongado, correspondiente a la transmisión de un millón de ráfagas. El tamaño de los paquetes IP sigue una distribución de Poisson centrada en 10 (slots), al igual que el tiempo de espera entre los mismos.

3.3 Variación del número de longitudes de onda

Este experimento tiene como objetivo validar que el aumento del número de longitudes de onda es inversamente proporcional al porcentaje de ráfagas descartadas. En la Fig. 7 se ven los resultados obtenidos para variaciones de 5 a 11 longitudes de onda. En este experimento el tamaño de la tabla de reservas se fijó en 350, el tamaño medio de los bursts en 8 (slots) y el tiempo de espera medio entre ráfagas en 10ns.

3.4 Variación del tamaño de las ráfagas

El objetivo de este experimento es el de validar que el aumento del tamaño de los paquetes IP es directamente proporcional al número de ráfagas descartadas, manteniendo constante el tamaño de la tabla. Para este experimento el tamaño de la tabla se fijó en 300, el número de longitudes de onda en 10, el tamaño medio de los paquetes 10 slots y el tiempo de espera medio entre ellos 10 ns.

Los resultados de este experimento se muestran en la Fig. 8, donde se puede comprobar que a partir de un cierto tamaño de las ráfagas, la tasa de descarte crece linealmente

3.6 Aleatorización de la posición de las ráfagas

Como se ha comentado al comienzo del artículo, gran parte de las pérdidas experimentadas en una red OBS o TSOBS son provocadas por las limitaciones de los mecanismos de reserva.

En los experimentos anteriores los nodo frontera rellenan la tabla de reservas se rellena de una forma secuencial y ordenada, lo que puede ocasionar que ráfagas entren en conflicto en los nodos troncales. Para evitar este problema, los nodos frontera pueden asignar una longitud de onda y un offset aleatorio a las ráfagas que crean (esto es, aleatorizar su posición en la tabla de reservas). En este experimento se ha tratado de evaluar si esta aleatorización en efecto mejora la tasa de descarte.

Se probaron varias arquitecturas de red, y donde se obtuvieron resultados más evidentes fue en una topología en forma de anillo (Fig. 10), donde se puede observar una mejora patente en la tasa de descarte al aleatorizar el offset y la longitud de onda escogidas (Fig. 9).

3 Conclusiones

En este artículo se ha presentado un modelo de red OBS completamente escrito en el lenguaje de descripción hardware VHDL. El objetivo de este modelo es servir de banco de pruebas para futuras implementaciones en FPGA de la sección de control de un nodo OBS, puesto que VHDL es en la actualidad la actualidad la opción predominante para trabajar con lógica programable.

Se han presentado las diversas simulaciones que se han realizado para validar el comportamiento del modelo, en las que se han variado tanto los parámetros de la red como su topología. Al realizar estas simulaciones se comprobó la influencia que tiene en las prestaciones de la red el algoritmo que se use para asignar a las ráfagas la longitud de onda y el offset en los nodos frontera. Se comprobó que si esta asignación se realiza de una manera aleatoria se consigue reducir la tasa de descarte de ráfagas.

Agradecimientos

Este trabajo ha sido financiado por la Fundación General de la Universidad Autónoma de Madrid, mediante los proyectos 016100 y 658006.

Referencias

[1] K. G. Coffman, A.M. Odlyzko. "The Size and Growth Rate of the Internet", 1998.

[2] Se-Yoon Oh. "A Data Burst Assembly Algorithm in Optical Burst Switching Networks". ETRI Journal, Vol. 24, Num. 4, agosto 2002.

[3] C. Qiao, M. Yoo. "Optical Burst Switching (OBS) – A New Paradigm for an Optical Internet". Journal of High Speed Networks, Num. 8, pp. 69 – 84, 1999.

[4] J. Ramamirtham, J. Turner. "Time Sliced Optical Burst Switching". Proc. IEEE INFOCOM 2003, San Francisco, abril 2003.

[5] W.A.Hanna. "Random Number Generator for VHDL Models". McDonnell Douglas Corp., 1993.

[6] J.S. Turner. "Terabit burst switching". Journal of High Speed Networks, No. 8, pp 3 – 16, 1999.

Criterio equitativo de asignación de longitudes de onda en redes SCWP de Conmutación Óptica de Paquetes

P. Pavón Mariño, F. J. González Castaño*, J. García Haro
Departamento de Tecnologías de la Información y las Comunicaciones, Universidad Politécnica de Cartagena
ETSI Telecomunicación, Campus Muralla de Mar s/n.
30202 – Cartagena (Murcia)

*Departamento de Ingeniería Telemática, Universidad de Vigo, ETSI Telecomunicación, Campus, 36310 – Vigo (Pontevedra)

Teléfono: 986 81 37 88 Fax: 986 81 21 16
E-mail: {pablo.pavon, joang.haro}@upct.es, javier@det.uvigo.es

Abstract. *The WASPNET project proposed a packet sequence criterion to preserve end-to-end packet order in Optical Packet Switching (OPS) networks controlled by the Scattered Wavelength Path (SCWP) operational mode. This paper shows that the WASPNET criterion inherently degrades performance, as a result of unbalanced wavelength usage in fiber links. As a solution, a round-robin sequence criterion is proposed. It eliminates non-uniform wavelength utilization in fiber links for any traffic pattern, as well as the performance impairments associated. We also propose an optimum scheduling algorithm for OPS output-buffered switch fabrics, which preserves packet order according to the new sequencing criterion.*

1 Introducción

1.1 Redes WDM de Conmutación Óptica de Paquetes

La Conmutación Óptica de Paquetes [1] (*Optical Packet Switching*, OPS) en redes de multiplexación por división en longitud de onda (*Wavelength Division Multiplexing*, WDM) es similar a la conmutación electrónica de paquetes tradicional, excepto que el campo de datos (*payload*) del paquete permanece en estado óptico, mientras que su cabecera se procesa electrónicamente. Este paradigma proporciona la mayor eficiencia de utilización de canal, al operar con la granularidad de los paquetes. Sin embargo, la función de conmutación óptica por paquetes y el ineludible almacenamiento óptico imponen fuertes exigencias a la tecnología fotónica. Por ello, aunque en diversos foros [1] se señala a la conmutación OPS como la solución definitiva para las redes WDM, no se espera disponer de una red troncal basada en dicha tecnología en un futuro inmediato.

La longitud de los paquetes ópticos representa aún una línea de discusión abierta. En este artículo asumiremos redes OPS síncronas, con ranuras temporales ajustadas al tamaño (fijo) de paquete. Esta alternativa proporciona mejores prestaciones, e implica la necesidad de etapas de sincronización óptica que alinean los paquetes a la entrada de los nodos. El proyecto europeo DAVID ha identificado este tipo de funcionamiento como la alternativa más prometedora para el troncal OPS, y ha sugerido un tamaño de paquete en el orden de 1 μs [2].

Se acepta generalmente que las conexiones de tráfico entre nodos OPS frontera se implementarán mediante caminos ópticos de paquete u *Optical Packet Paths (OPP)*. Durante la provisión de un OPP, se fija la secuencia de fibras que atravesará el tráfico de esa conexión, desde el nodo de entrada hasta el de salida. La ventaja de la conmutación OPS estriba en que (i) la capacidad de cada OPP puede ser mucho menor que la capacidad de un longitud de onda, (ii) el número de OPPs que atraviesan un enlace puede ser mucho mayor que el de longitudes de onda. El modo de operación de una red OPS [3] establece el criterio de correspondencia entre los OPPs y las longitudes de onda en cada salto. Según el modo de operación SCWP (*Scattered Wavelength Path*), la longitud de onda de los paquetes en cada salto puede variar. Los paquetes que acceden a un nodo demandan la fibra de salida que les corresponde (en función del OPP al que pertenecen), pero el conmutador puede elegir dinámicamente su longitud de onda de salida. Este grado de libertad adicional permite a los planificadores tomar una decisión conjunta sobre retardo y longitud de onda de salida, potenciando el efecto de multiplexado estadístico. Debido a la mejora de prestaciones resultante, el modo de operación SCWP podría ser el mecanismo natural en las futuras redes OPS [3].

1.2 Secuencia de paquetes en redes OPS

En una red troncal OPS, es necesario preservar la secuencia de paquetes extremo a extremo, para evitar el coste de reordenación en el nodo de salida (etapas electrónicas con memorias muy grandes, debido a la alta velocidad de las líneas ópticas). Por tanto, la red OPS debe mantener el orden de los paquetes en cada

salto. Esto requiere que cada nodo conozca el orden entre los paquetes entrantes, a fin de mantenerlo en la asignación de retardos y longitudes de onda de salida. Preferiblemente, esto se debe conseguir sin recurrir a un campo contador adicional en la cabecera de los paquetes ópticos. Opción poco deseable, debido a la necesidad de reducir al máximo el tamaño de la cabecera óptica. La alternativa consiste en diseñar un criterio de orden de paquetes que se base en su instante de llegada y longitud de onda. El instante de llegada no basta, porque se pueden recibir simultáneamente distintos paquetes de un mismo OPP en longitudes de onda diferentes de una misma fibra.

Hasta donde alcanza el conocimiento de los autores, el único criterio de secuencia independiente de cabeceras para redes SCWP síncronas se propuso en el proyecto WASPNET [3]. Éste consiste en transmitir paquetes simultáneos en longitudes de onda consecutivas, comenzando siempre por la longitud de onda más baja λ_0. Chia *et al*. [4] presentaron un algoritmo de planificación para el conmutador WASPNET realimentado [3] que satisface dicho criterio. En [5], los autores propusieron un algoritmo de planificación para conmutadores OPS capaz de emular un comportamiento de colas a la salida. Su principal ventaja reside en que ofrece prestaciones óptimas para una complejidad razonable, preservando el orden entre paquetes según el criterio WASPNET. Denotaremos este algoritmo como *planificador SCWP básico*.

El resto de este artículo está organizado como sigue: en la sección 2 se identifican y evalúan los problemas de la aplicación del criterio WASPNET. La sección 3 propone un nuevo criterio de secuencia de paquetes que solventa los problemas identificados. La sección 4 adapta el planificador SCWP básico al nuevo criterio de secuencia. La sección 5 evalúa la mejora de prestaciones resultante. Finalmente, la sección 6 recoge las principales conclusiones.

2 Selección no uniforme de longitudes de onda

La Figura 1(a) muestra cómo el criterio de secuencia WASPNET causa un desequilibrio en la asignación de longitudes de onda. Ello se debe a que las longitudes de onda bajas se utilizan más frecuentemente que las altas. En cada ranura temporal, si λ_i está ocupada, λ_{i-1} también lo está, mientras que lo contrario no se cumple. En consecuencia, para cualquier fibra de la red, la utilización media x_i de la longitud de onda i=0...n-1 siempre cumple $1 \geq x_0 \geq x_1 \geq ... \geq x_{n-2} \geq x_{n-1} \geq 0$.

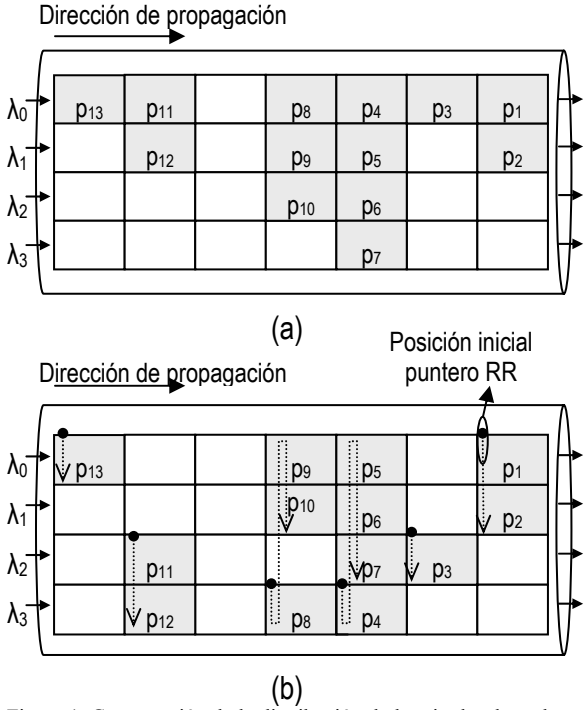


Figura 1. Comparación de la distribución de longitudes de onda en una fibra (cuatro longitudes de onda, λ_0,..., λ_3). (a) Criterio WASPNET, (b) criterio de secuencia *round-robin*. El símbolo p_i denota un paquete.

Para la evaluación del desequilibrio de utilización de las longitudes de onda en las fibras, simulamos la red OPS de la figura 2. La parte izquierda de la figura representa un conjunto de N nodos OPS con N fibras de entrada/salida y n longitudes de onda por fibra. El tráfico inyectado en el punto (X) de cada fibra de entrada es la agregación de n fuentes Bernoulli uniformes de parámetro ρ, $0 \leq \rho \leq 1$. La fibra de salida 0 de cada nodo izquierdo se conecta al nodo OPS A, y la fibra de salida 1 se conecta al nodo OPS B. Los nodos del lado izquierdo de la figura son conmutadores KEOPS $nN \times nN$ [6], capaces de emular el comportamiento de colas a la salida [5].

Para evaluar el efecto de la asignación no uniforme de longitudes de onda que resulta de la aplicación del criterio de orden WASPNET, aplicamos el planificador SCWP básico de [5] a los conmutadores KEOPS y observamos el punto (Y) de la figura 2. Estimamos por simulación las probabilidades de selección de longitud de onda (método *Batch Means* [7], intervalos de confianza de 99%, tolerancia de 1%, límite de $5 \cdot 10^{10}$ paquetes), bajo tráfico de entrada Bernoulli uniforme de carga media ρ, y parámetros $N \in \{2,4,8\}$ y $n \in \{2,4,8,16,32,64\}$. Ocho fibras de entrada/salida son suficientemente representativas para evaluar topologías troncales WDM, en las que los nodos tienen típicamente un grado de conectividad menor que cinco. El número de retardos del conmutador se dimensiona para que las pérdidas de paquetes sean despreciables en cualquier selección (ρ, N, n). Como medida del desequilibrio de selección de longitudes de onda, la figura 3 muestra la relación entre la mayor utilización (λ_0) y la menor

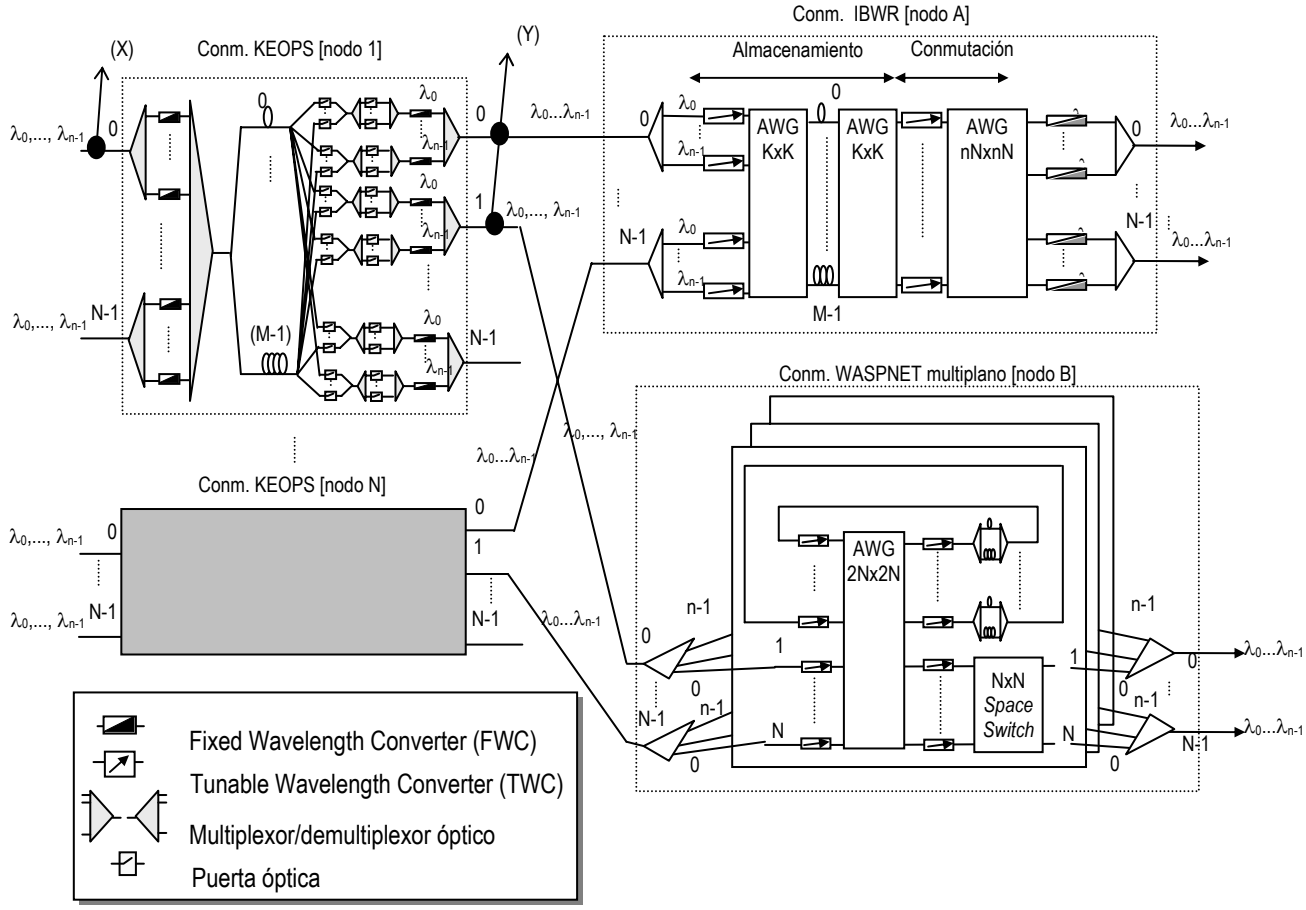


Figura 2. Diagrama de la red SCWP de referencia.

(λ_{n-1}), para cada tupla (ρ, N, n). Se puede observar que el desequilibrio es mayor para cargas bajas y medias y crece con n. Además, se aprecia un pequeño aumento de la uniformidad de la selección para valores altos de N. Nótese que los valores pueden ser muy elevados para cargas típicas de red (como ρ=0.5).

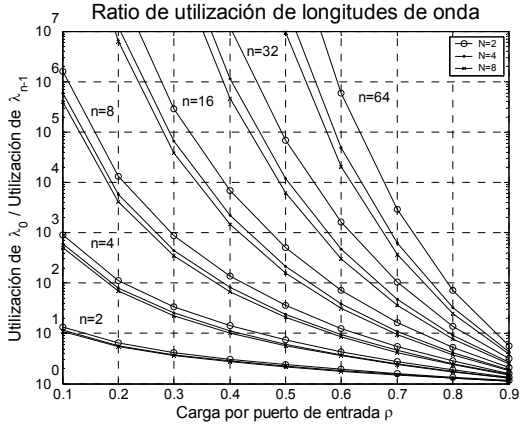


Figura 3. Relación de utilización de longitudes de onda: utilización (λ_0) / utilización (λ_{n-1}), en las fibras de salida de un conmutador OPS con colas a la salida, en una red que sigue el criterio WASPNET.

En este punto de nuestro análisis, podría argumentarse que una distribución no uniforme de longitudes de onda puede no ser gravosa, puesto que no afecta ni a la pérdida de paquetes ni al retardo en un conmutador con colas a la salida. Esto se debe a la etapa de conmutación sin bloqueo que está presente a la entrada de dichas arquitecturas, y que hace que la distribución de longitudes de onda de entrada sea irrelevante para la asignación de retardos y longitudes de onda de salida. Como ejemplos de arquitecturas de conmutación con esta propiedad podemos citar el conmutador WASPNET monoplano [3] y las arquitecturas OPS de colas a la salida (como KEOPS, figura 2), *Space Switch*, o el conmutador OBWR (*Output-Buffered Wavelength-Routed*) [5].

Sin embargo, también hemos identificado una notable degradación de prestaciones en otras arquitecturas OPS, en las que una distribución no uniforme de longitudes de onda en el tráfico de entrada (como la que resulta del criterio WASPNET) causa mayores retardos de paquete. Esta dependencia con la distribución de longitudes de onda de entrada se debe a la ausencia de una primera etapa de conmutación sin bloqueo. Los nodos a la derecha de la figura 2 representan dos arquitecturas en las que hemos observado esta merma de prestaciones: el conmutador IBWR (*Input-Buffered Wavelength-Routed*) [8], y el conmutador WASPNET multiplano [3][4]. Los resultados de las evaluaciones realizadas se presentan y discuten en la sección 5.

3 Criterio de secuencia *round-robin*

Esta sección describe el criterio de secuencia *round-robin* que se propone en este artículo. Sus principales propiedades son:

- No necesita almacenar información de secuencia en la cabecera de los paquetes.
- Garantiza una utilización uniforme de longitudes de onda en todas las fibras de la red, independientemente del patrón de tráfico.

Sean p_i y p_{i+1} paquetes ordenados y consecutivos que se transmiten por un enlace WDM en las ranuras temporales $t(p_i)$ y $t(p_{i+1})$ y en las longitudes de onda $\lambda(p_i)$ y $\lambda(p_{i+1})$, respectivamente. El criterio *round-robin* especifica que: (i) $t(p_{i+1}) \geq t(p_i)$, y (ii) $\lambda(p_{i+1})=(\lambda(p_i)+1) \bmod n$, donde n es el número de longitudes de onda de la fibra y (a mod b) es el resto de a/b para dos enteros a y b cualesquiera.

El resultado, como se muestra en la figura 1(b), es una dispersión *round-robin* exacta a través de las longitudes de onda, para cualquier patrón de tráfico en la fibra. Si k_i es la suma total de los paquetes que se transmiten en la longitud de onda λ_i, i=0,...,n-1 de la fibra a lo largo de T ranuras, se cumple que $|k_i-k_j| \in \{0,1\}$, $\forall$ i,j=0,...,n-1, $\forall T$>0. En otras palabras: en cualquier intervalo de tiempo, el número de paquetes que se transmiten en dos longitudes de onda diferentes cualesquiera difiere, a lo sumo, en uno.

Consideremos ahora los cambios que supondría la adopción del criterio *round-robin* en el diseño de la red OPS. Una primera consecuencia es la necesaria modificación de los algoritmos de planificación SCWP de las arquitecturas OPS, para satisfacer las nuevas especificaciones de ordenamiento. Como se puede deducir de la figura 1(b), el criterio requiere que cada nodo "recuerde" la longitud de onda del último paquete de la secuencia recibido/transmitido, a través de ranuras temporales sucesivas. Una implementación trivial de esta funcionalidad implica un conjunto de punteros *round-robin* para seguir la secuencia de los paquetes:

1) Un puntero *round-robin* por fibra de entrada, que apunta a la longitud de onda del siguiente paquete en la secuencia de entrada. Cuando aparece un nuevo paquete en esta longitud de onda, se incrementa el puntero en modo *round-robin*.

2) Un puntero *round-robin* por fibra de salida, para determinar la longitud de onda de salida del siguiente paquete a transmitir. Cuando se transmite un nuevo paquete, el puntero se incrementa en modo *round-robin*.

Este método también precisa la sincronización de los punteros al iniciar los sistemas: el puntero de cada fibra de entrada de un nodo se debe sincronizar con el puntero de la fibra de salida del nodo previo. La implementación del criterio *round-robin* no es inmediata y se debe estudiar independientemente para cada algoritmo de planificación. La sección 4 presenta el planificador SCWP uniforme para arquitecturas OPS con colas a la salida, que satisface el nuevo criterio de secuencia.

4 Planificador SCWP uniforme

En un trabajo previo [5], los autores propusieron el *planificador SCWP básico* para arquitecturas OPS con colas a la salida. Este planificador tiene prestaciones óptimas, y preserva la secuencia de paquetes según el criterio WASPNET. Nuestro objetivo ha sido modificar el algoritmo que se propuso en [5] para cumplir el nuevo criterio *round-robin*, manteniendo a un tiempo las prestaciones óptimas. Esto se consigue con el algoritmo de la figura 4:

Planificador SCWP uniforme

```
 1.  for fiberCounter = 0 to N-1 do
 2.    f_in = (f_0 + fiberCounter) mod N
 3.    for wavCounter = 0 to n-1 do
 4.      if (packet p in input (f_in , λ_in [f_in]) then
 5.        f_out = output fiber p ( = opp (p) )
 6.        if (delay [f_out] < M) then
 7.          associate delay [f_out] to packet p
 8.          lastDelayOccup [f_out] ++
 9.          if (lastDelayOccup [f_out] == n) then
10.            lastDelayOccup [f_out] = 0
11.            delay [f_out] ++
12.          endif
13.          associate λ_out [f_out] to packet p
14.          λ_out [f_out] = (λ_out [f_out] + 1) mod n
15.        endif
16.        λ_in [f_in] = (λ_in [f_in] + 1) mod n
17.      else
18.        /* para cada f_in, las λ se seleccionan
19.        hasta encontrar una vacía, o hasta que se
20.        han seleccionado n */
21.        break;
22.      end
23.    endfor
24.  endfor
25.
26.  /* Realizado al final de cada ranura temporal */
27.  f_0 = (f_0 + 1) mod N  /* garantiza equidad */
28.  for f_out = 0 to N-1 do
29.    if (delay [f_out] == 0)
30.      lastDelayOccup [f_out] = 0;
31.    else
32.      delay [f_out] --;
33.    endif
34.  endfor
```

Figura 4. Planificador SCWP uniforme.

Descripción del algoritmo:

- Exploración de puertos de entrada. Para mantener el orden de los paquetes, cada nodo necesita un puntero de exploración de longitudes de onda por fibra de entrada ($\lambda_{in}[f_{in}]$) que sigue la secuencia de los paquetes. Para considerar todas las fuentes de tráfico (fibras de entrada) equitativamente, el algoritmo rota el índice de la primera fibra de entrada que se comprueba en cada ranura (f_0).
- Asignación de retardo y longitud de onda de salida. La variable $delay[f_{out}]$ almacena lo que

llamamos "retardo activo" de la fibra de salida f_{out}: el retardo que se está asignando en ese momento a los paquetes destinados a f_{out}. La variable `lastDelayOccup` almacena el número de paquetes a los que se asigna el retardo activo. Se alcanzan prestaciones óptimas puesto que (i) cualquier paquete obtiene el retardo más corto disponible (retardo activo, línea 7), (ii) cuando se asignan n_{out} paquetes, el algoritmo selecciona el siguiente retardo (líneas 9-11), (iii) a consecuencia de ello, sólo se pierde un paquete destinado a la fibra de salida f_{out} (línea 6) cuando todos los M retardos de esa fibra de salida tienen n_{out} paquetes. El algoritmo satisface el criterio de secuencia, puesto que los paquetes consecutivos que se transmiten por la fibra de salida f_{out} obtienen sus longitudes de onda de salida (líneas 13-14) a partir del puntero *round-robin* $\lambda_{out}[f_{out}]$.

5 Evaluación

Para estimar la degradación de prestaciones debida al sesgo de la distribución de las longitudes de onda de entrada, comparamos el retardo medio sufrido por los paquetes en los conmutadores IBWR y WASPNET multiplano (figura 2), en dos escenarios (I) y (II):

(I). Aplicar el criterio de secuencia WASPNET a la red. Ello implica implementar el planificador SCWP básico en los conmutadores KEOPS del lado izquierdo de la figura 2, lo que causa un desequilibrio en la distribución de longitudes de onda en el tráfico que accede a los nodos A y B.

(II). Aplicar el criterio de secuencia *round-robin* que se propone en la sección 3. Ello conlleva planificar los conmutadores KEOPS según el algoritmo SCWP uniforme propuesto en la sección 4. Esto implica una utilización uniforme de longitudes de onda en el tráfico que accede a los nodos A y B (y, lógicamente, en todas las fibras de la red, si el criterio se aplica de manera generalizada).

En ambos escenarios, el planificador empleado para la arquitectura WASPNET es el propuesto en [4], y el empleado para la arquitectura IBWR el propuesto en [9].

Sean $D_{nw}(n,N)$ y $D_{uw}(n,N)$ los retardos por paquete (en número de ranuras temporales) en un nodo A o B que recibe tráfico de fuentes con distribución de longitudes de onda no uniforme o uniforme, respectivamente. Sea $L(n,N)=D_{nw}(n,N)/D_{uw}(n,N)$ el incremento relativo de retardo de paquete en un nodo A o B (es decir, tomando como referencia la distribución uniforme de longitudes de onda). Las figuras 5(a) y 5(b) muestran $\log_2(L(n,N))$ para cargas de entrada de 50% y 80% en una arquitectura IBWR, y las figuras 5(c) y 5(d) muestran $\log_2(L(n,N))$ para las mismas cargas de entrada en una arquitectura WASPNET multiplano. Nótese que $L(n,N)>1.0\ \forall n,N$.

Los valores de $L(n,N)$ pueden ser tan altos como $L(32,8)=8\cdot10^3$ para ρ=50% en el nodo A, $L(64,2)=21.14$ para ρ=80% en el nodo A, $L(32,4)=1.8\cdot10^3$ para ρ=50% en el nodo B o $L(64,2)=381.42$ para ρ=80% en el nodo B. Estos resultados revelan en impacto de una distribución no uniforme de longitudes de onda de entrada, en las dos arquitecturas escogidas como ejemplo.

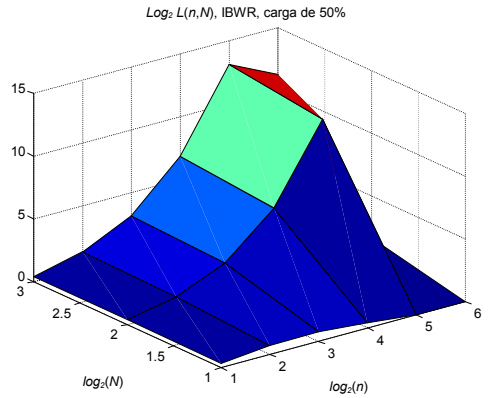


Figura 5(a). $\log_2(L(n,N))$, ρ=50%, nodo A (IBWR).

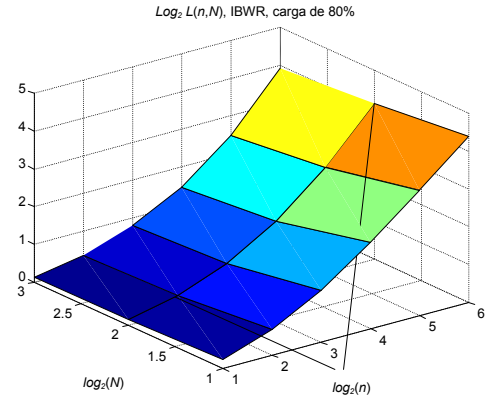


Figura 5(b). $\log_2(L(n,N))$, ρ=80%, nodo A (IBWR).

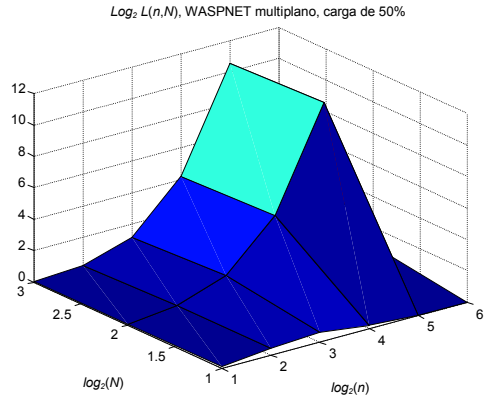


Figura 5(c). $\log_2(L(n,N))$, ρ=50%, nodo B (WASPNET multiplano).

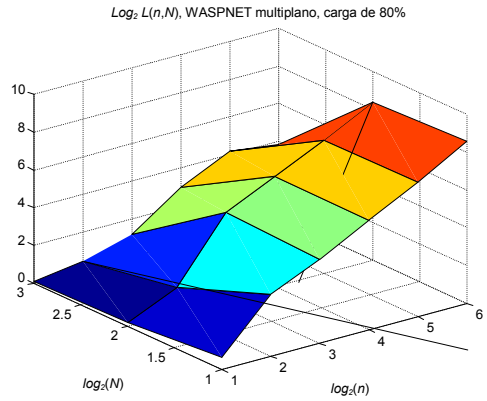


Figura 5(d). $\log_2(L(n,N))$, ρ=80%, nodo B (WASPNET multiplano).

3 Conclusiones

Este artículo analiza los problemas que causa el criterio WASPNET para el mantenimiento del orden de paquetes extremo a extremo en una red OPS SCWP. Estos problemas se deben a la utilización no equilibrada de longitudes de onda, que es una consecuencia intrínseca del criterio WASPNET. Como solución, se propone un nuevo criterio de secuencia, al que llamamos criterio *round-robin*. Dicho criterio mantiene la secuencia de paquetes extremo a extremo sin necesidad de un campo específico en su cabecera, y garantiza una utilización uniforme de longitudes de onda en todas las fibras de la red, bajo cualquier patrón de tráfico. Otra contribución de este artículo es la propuesta de un planificador SCWP uniforme para arquitecturas OPS con capacidad de emular colas a la salida. El nuevo algoritmo proporciona prestaciones óptimas, manteniendo la secuencia de paquetes según el nuevo criterio *round-robin*.

Agradecimientos

Este trabajo se ha financiado con los proyectos nacionales TEC2004-05622-C04-01/TCM (CAPITAL) y TEC2004-05622-C04-02/TCM (ARPaq), y con el proyecto PGIDIT04TIC322003PR de la Xunta de Galicia.

Referencias

[1] M. O'Mahony, D. Simeonidou, D. Hunter, A. Tzanakaki. "The application of optical packet switching in future communication networks". IEEE Communications Magazine, pp. 128-135, vol. 39, no. 3, ISSN: 0162-6804 (2001).

[2] L. Dittman *et al.,* "The European IST Project DAVID: A Viable Approach Toward Optical Packet Switching". IEEE Journal on Selected Areas in Communications, pp. 1026-1040, vol. 21, no. 7, ISSN: 0733-8716 (2003).

[3] D. Hunter *et al.* "WASPNET: A Wavelength Switched Packet Network". IEEE Communications Magazine, pp. 120-129, vol. 37, no. 3, ISSN: 0162-6804 (1999).

[4] M. Chia, D. Hunter, I. Andonovic, P. Ball, I. Wright, S. Ferguson, K. Guild, M. O'Mahony. "Packet loss and delay performance of feedback and feed-forward arrayed-waveguide gratings-based optical packet switches with WDM inputs-outputs". IEEE Journal of. Lightwave Technology, pp. 1241-1254, vol. 19, no. 9, ISSN: 0733-8724 (2001).

[5] P. Pavón Mariño, J. García Haro, J. Malgosa Sanahuja, F. Cerdán. "Scattered Versus Shared Wavelength Path Operation, Application to Output Buffered Optical Packet Switches. A Comparative Study". SPIE/Kluwer Optical Networks Magazine, pp. 134-145, vol. 4, no. 6, ISSN: 1388-6916 (2003).

[6] C. Guillemot *et al.* "Transparent optical packet switching: the European ACTS KEOPS project approach". IEEE Journal of Lightwave Technology, pp. 2117-2134, vol. 16, no. 12, ISSN: 0733-8724 (1998).

[7] A. M. Law, J. S. Carson. "A sequential procedure for determining the length of a steady-state simulation". Operations Research, pp. 1011-1025, vol. 27, ISSN: 0030-364X (1979).

[8] W. Zhong, R. Tucker. "Wavelength routing-based photonic packet buffers and their applications in photonic packet switching systems". IEEE Journal of Lightwave Technology, pp. 1737-1745, vol. 16, no. 10, ISSN: 0733-8724 (1998).

[9] P. Pavón Mariño, J. García Haro, J. Malgosa Sanahuja, F. Cerdán. "Maximal Matching Characterization of Optical Packet Input-Buffered Wavelength Routed Switches". En Proc. of 2003 IEEE Workshop on High Performance Switching and Routing HPSR 2003. Torino (Italia), Junio 2003, pp. 55-60, ISBN: 0-7803-7710-9.

Modelado estocástico de TCP sobre redes OBS*

Juan Ramón Troncoso Pastoriza y Manuel Fernández Veiga
Departamento de Ingeniería Telemática. Universidad de Vigo
ETSI de Telecomunicación. C/ Maxwell S/N. Campus Universitario.
36310 - Vigo (Pontevedra)
E-mail: jramon@det.uvigo.es

Abstract *In this article, a new analytical description of the send rate of the Transmission Control Protocol (TCP) over an Optical Burst Switching (OBS) network is formulated, to study the impact that the future use of this switching technology in the backbone networks of Internet can have on the performance of such protocol. It is divided in two parts: first, modeling the drop probability of an OBS network, as its behavior differs from the traditional electronic packet switching networks, and second, modeling the TCP throughput over an OBS network. Results in both models are contrasted with ns-2 simulations, and conclusions about differences in low-rate and high-rate flows are extracted.*

1. Introducción

OBS [2] (*Optical Burst Switching*) es una tecnología que integra la conmutación de paquetes y la transmisión WDM (*Wavelength Division Multiplexing*) con la intención de aumentar el aprovechamiento del ancho de banda de las fibras ópticas. En una red OBS, la unidad básica de datos es la ráfaga —una agrupación de paquetes consecutivos— que se conmuta por medios puramente ópticos, sin conversión previa al dominio electrónico, y además sin posibilidad de espera, dada la inexistencia material de dispositivos ópticos de memoria. Así, un pequeño paquete de control precede a cada ráfaga para que los elementos internos de conmutación de un nodo estén preparados en el instante en que llegue. Aunque el flujo de paquetes de control se procesa electrónicamente, este tiempo se amortiza entre todos los paquetes de la ráfaga, por lo que la velocidad de conmutación se aproxima a la de transmisión. Una red OBS es, por tanto, una red de conmutación carente de *buffers*, a cuyos nodos están conectados *routers* electrónicos convencionales que se llamarán de ingreso si introducen tráfico en la subred óptica y de egreso si allí finaliza la conmutación en el dominio óptico.

TCP [7] ha sido el protocolo de transporte por excelencia para las aplicaciones elásticas durante varias décadas, proporcionándoles un servicio orientado a conexión, fiable y ordenado; versión tras versión se ha ido adaptando a una red en la que la principal causa de pérdida de paquetes es la congestión. Esta es una característica que diferencia a una red de conmutación de paquetes electrónica (EPS) de una red OBS, en la que la contienda puede llegar a ser tanto o más importante que la congestión en su contribución a la tasa de pérdida.

En los últimos años se han venido desarrollando una serie de modelos analíticos [1, 5, 6] para predecir el comportamiento de TCP, pero las suposiciones realizadas no se adecuan al comportamiento de una red OBS. Dado que una comprensión del comportamiento dinámico de TCP sobre este tipo de redes resulta esencial para utilizar eficazmente el potencial de la conmutación óptica de ráfagas, en este artículo se desarrolla un nuevo modelo analítico que contempla el efecto de la técnica OBS en el rendimiento del protocolo. En la sección 2 se modela la probabilidad de descarte en una red OBS, diferenciando cada uno de sus elementos; en la sección 3 se introduce el nuevo modelo de TCP. La validez de ambos modelos se ha comprobado mediante simulaciones realizadas con *ns-2*. Las conclusiones se muestran en la sección 4.

2. Probabilidad de pérdidas

Con objeto de alcanzar una fórmula para la probabilidad de pérdida de paquete en una red OBS, se adoptarán las siguientes hipótesis:

a) Se diferenciará entre nodos de ingreso, en los que se realiza el ensamblado de paquetes en ráfagas, y nodos internos de la red, que sólo conmutan ráfagas.

b) No se considera la posible existencia de retardos en los nodos (*Fiber Delay Lines* FDLs).

c) Las llegadas de paquetes a los nodos de ingreso se producen según un proceso de Poisson de tasa λ [paquetes/s], y todos los paquetes son del

*Este trabajo ha sido subvencionado por el Ministerio de Educación y Ciencia dentro del proyecyo TIC2003-09042-C03-03 del Plan Nacional de I+D+I (parcialmente financiado con fondos FEDER) y por la Dirección Xeral da Xunta de Galicia con el proyecto PGIDT04PXIC32203PN (Plan Galego de Investigación, Desenvolvemento e Innovación Tecnolóxica).

mismo tamaño. Asimismo, se considera [3] que las llegadas de ráfagas a los nodos internos también siguen un proceso de Poisson de tasa λ_r [ráfagas/s].

d) La estrategia de ensamblado de paquetes en ráfagas está basada en tiempos, lo que permitirá diferenciar más adelante entre flujos rápidos, lentos y de tasa media. Se denotará por T_s el tiempo de ensamblado de ráfaga.

e) Se considerará el comportamiento de los nodos independiente entre sí, para que el problema sea tratable analíticamente.

2.1. Nodos de ingreso

Los nodos de ingreso realizan la conversión electro-óptica de la señal de información, así como el ensamblado de los paquetes en ráfagas.

Para describir su comportamiento, se considerará que se ensamblan juntos en una misma ráfaga todos los paquetes que llegan durante un intervalo de tiempo de duración T_s, que se inicia con la llegada del primer paquete, y que siguen la misma ruta. Tras el intervalo de ensamblado se vuelve a un período de silencio a la espera de un nuevo paquete. En el dominio eléctrico, para cada enlace de salida estos nodos deben poseer *buffers* (de tamaño $N_{\max}$, que se considerará el límite al tamaño de ráfaga) para el almacenamiento temporal de los paquetes que llegan.

Si durante un intervalo T_s llegan más de $N_{\max}$ paquetes (contando el paquete que inició la temporización), serán descartados. Es razonable suponer que el cuello de botella en estos nodos es el límite al tamaño de ráfaga, y no el tiempo de transmisión a través del enlace de salida. Así, se despreciará dicho tiempo de transmisión, y se tomará el límite al tamaño de ráfaga como única causa de descarte.

Dado que el proceso de llegadas de paquetes a los nodos de ingreso se ha supuesto poissoniano de tasa λ, la probabilidad de descarte (p_{dpna}) se puede calcular como la proporción media de paquetes descartados en un intervalo. Se denotará por $A = \lambda \cdot T_s$ al número medio de llegadas en un intervalo, por lo que, en media, se tendrán $A+1$ paquetes.

El número medio de descartes viene dado por el número de paquetes que llegan al nodo por encima de $N_{\max} - 1$ en un intervalo T_s (1). La probabilidad de descarte se puede obtener dividiendo $\overline{N_d}$ entre el número medio de paquetes que llegan en un intervalo de ráfaga $(A+1)$:

$$p_{\text{dpna}}(A, N_{\max}) = \frac{\overline{N_d}}{A+1}. \qquad (2)$$

Se ha supuesto que en el momento en que comienza la transmisión de la ráfaga, los *buffers* ya están disponibles para almacenar nuevos paquetes, suposición razonable tanto ante tasas de tráfico bajas (sin problemas de almacenamiento) como elevadas (descarte despreciable frente al producido en T_s).

Por último, si se supone que todos los nodos de ingreso de la red tienen una carga similar, y que el tamaño máximo de ráfaga ($N_{\max}$) es una constante de la red, se puede obtener la distribución del tamaño de las ráfagas que circulan por la misma.

Dado que no comienza el ensamblado de una ráfaga hasta la llegada de un paquete, el tamaño de ráfaga estará comprendido entre 1 y $N_{\max}$, y tendrá una distribución poissoniana limitada por $N_{\max}$; el tamaño medio de ráfaga se muestra en la ecuación (3). En la práctica, si $N_{\max} \geq 10 \cdot A$ y $A \geq 1$, se puede considerar $\overline{N} = A + 1$, despreciando la probabilidad de descarte en los nodos de ingreso.

2.2. Nodos internos

Estos nodos conmutan ráfagas. En esta parte de la red se pueden hacer varias suposiciones en cuanto a la contienda; en el caso más sencillo, los nodos descartan ráfagas completas si están ocupados; de este modo, la ráfaga que se encontraba en el conmutador no se pierde, pues los recursos de conmutación estaban reservados de antemano.

Con esta suposición, la probabilidad de descarte de una ráfaga no depende de su tamaño, sino solamente del instante de llegada. Dada la hipótesis de llegadas poissonianas (λ_r), se puede modelar cada nodo como un sistema de colas $M/G/1/1$ [4], para el que tiene validez la primera fórmula de Erlang en el cálculo de la probabilidad de descarte, debido a su independencia con la distribución del tiempo de servicio demandado. Particularizándola para este caso ($m = 1$), y definiendo C [paquetes/s] como la capacidad del enlace de salida, y $A_r = \lambda_r \cdot \overline{N}/C$ como la intensidad de tráfico de entrada al nodo, se obtiene:

$$p_{\text{drni}}(A_r) = \frac{A_r^m}{m! \sum_{k=0}^{m} \frac{A_r^k}{k!}} = \frac{A_r}{1 + A_r}.$$

Para hallar la probabilidad de descarte de paquete basta recordar la hipótesis según la cual el descarte de una ráfaga es independiente de su tamaño, por lo que $p_{\text{dpni}}(A_r) = p_{\text{drni}}(A_r)$.

En esta sección se ha supuesto un esquema de señalización a una vía (*one-way reservation*), de

$$\overline{N_d} = \sum_{k=N_{\max}}^{\infty} (k - N_{\max} + 1) \cdot P_k = A - N_{\max} + 1 + e^{-A} \cdot \left[(N_{\max} - 1) \cdot \sum_{k=0}^{N_{\max}-1} \frac{A^k}{k!} - A \cdot \sum_{k=0}^{N_{\max}-2} \frac{A^k}{k!} \right] \qquad (1)$$

$$\overline{N} = e^{-A} \sum_{k=1}^{N_{\max}-1} k \frac{A^{k-1}}{(k-1)!} + N_{\max} \cdot \left(1 - e^{-A} \sum_{k=0}^{N_{\max}-2} \frac{A^k}{k!} \right) \qquad (3)$$

modo que los descartes se producen efectivamente en los nodos internos. Si el esquema fuese extremo a extremo (*end-to-end reservation*), las ráfagas serían descartadas directamente en el nodo de ingreso, no habría descartes en los nodos internos, y como contrapartida la latencia sería mayor. Pero dada la hipótesis de independencia entre nodos, también se podría aplicar el resultado anterior de forma aproximada a redes con reserva extremo a extremo.

2.3. Nodos de egreso

Los nodos de egreso desensamblan los paquetes de las ráfagas y los devuelven al dominio eléctrico. Estos nodos encaminan paquetes, no ráfagas, así que el cálculo de la probabilidad de pérdida es bastante más complicado que en los nodos internos, pues los paquetes de una misma ráfaga pueden tener distintos destinos.

No obstante, estos nodos se pueden dividir conceptualmente (y también físicamente) en dos partes:

a) Una parte de conversión óptica-eléctrica, desensamblado y demultiplexado de los paquetes de las ráfagas hacia los enlaces de salida. Para esta sección del nodo se puede usar la misma fórmula aplicable a los nodos internos, sustituyendo la capacidad del enlace de salida por la velocidad de procesamiento. Para ocupaciones moderadas de la red, esta probabilidad de descarte se podría despreciar, al igual que se ha despreciado el retardo de procesado en los nodos de ingreso.

b) La segunda parte corresponde al encolado y transmisión eléctrica, que se puede asimilar a la zona exterior a la red OBS, y que por lo tanto no se tendrá en cuenta en este análisis.

2.4. Modelo completo

A partir de la topología de la red, y manteniendo la hipótesis de independencia entre nodos, se puede hallar analíticamente la probabilidad de descarte de paquete en el segmento OBS de una ruta determinada, según la ecuación (4); donde λ representa la tasa de llegadas de paquetes que siguen esta ruta al nodo de ingreso; T_s es el tiempo de ensamblado de ráfagas en el nodo de ingreso; $N_{\max}$ es el tamaño máximo de ráfaga en número de paquetes; $\overline{N}$ es el tamaño medio de ráfaga, calculado según la ecuación (3); $\overrightarrow{\lambda_r}$ es el vector de tasas medias (totales) de llegadas de ráfagas a cada uno de los m nodos intermedios; y $\overrightarrow{C}$ es el vector de capacidades de los enlaces atravesados.

2.5. Comprobación experimental

En las pruebas realizadas se ha utilizado una red OBS con un esquema de ensamblado de ráfagas basado en tiempos, de umbral 10 ms, planificadores LAUC-VF [2] (*Latest Available Unscheduled Channel with Void Filling*) con un solo hueco, y protocolo de reserva JET [2] (*Just-Enough-Time*), con un tiempo de conmutación de 1 μs. Para evitar la influencia de este tiempo, se han utilizado paquetes de tamaño fijo suficientemente grande (1 MB).

En la Fig. 1 se muestran los resultados obtenidos para una topología con dos nodos internos dispuestos en serie y N nodos de ingreso conectados al primero con fuentes de tráfico poissoniano. Para que el tráfico de ráfagas que accede al segundo también sea poissoniano, se tienen M fuentes cuyo tráfico atraviesa solamente dicho nodo. Se monitorizan las pérdidas producidas en las N primeras fuentes.

Como se puede observar en la Fig. 1, se han realizado pruebas para varios tamaños de *buffer* de los nodos de ingreso, con $N = 40$ y $M = 20$; en el caso de *buffers* pequeños, la probabilidad de descarte está dominada por las pérdidas en los nodos de ingreso, y el modelo es muy preciso. En el caso de tamaños de *buffer* mayores, los descartes se producen fundamentalmente en el interior de la red OBS, y, dado que el tráfico se suaviza a su paso por cada nodo, la probabilidad de descarte dada por el modelo es ligeramente pesimista. Para probabilidades de descarte bajas ($p_{\mathrm{dp}} < 0{,}1$), el modelo se comporta bien, mientras que para proporciones de descarte mayores, se puede tomar como una cota superior, aunque en topologías más complejas es probable que la hipótesis de llegadas poissonianas a los nodos internos sea más precisa.

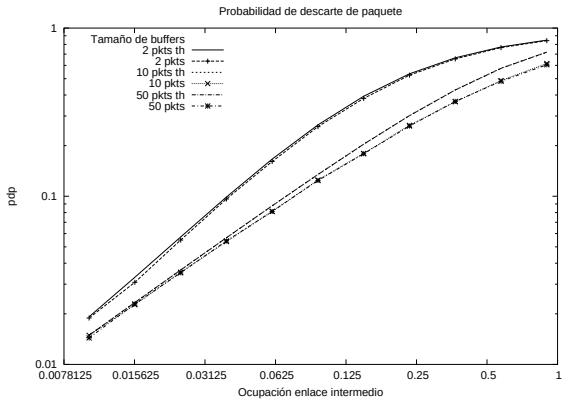


Figura 1: p_{dp} en una topología sencilla.

$$p_{\mathrm{dp}}(\lambda, T_s, N_{\max}, \overline{N}, \overrightarrow{\lambda_r}, \overrightarrow{C}) = 1 - (1 - p_{\mathrm{dpna}}(\lambda \cdot T_s, N_{\max})) \prod_{k=0}^{m-1} \left(1 - p_{\mathrm{dpni}}(\lambda_{\mathrm{rk}} \cdot \overline{N}/C_k)\right) \qquad (4)$$

3. Modelo de TCP

En esta sección se desarrolla un nuevo modelo de la tasa de envío de un flujo TCP al atravesar una red OBS. Se tomará el trabajo realizado en [5] como punto de partida, dado que éste es uno de los modelos más eficientes, precisos y sencillos de entre los existentes para TCP. Se utilizará un método similar y, en lo posible, la misma nomenclatura introducida en dicho artículo.

Para facilitar el tratamiento analítico del problema se hará una división entre flujos rápidos, lentos y de tasa media, como la introducida en [8], en función de la tasa máxima de transmisión (λ), limitada por el enlace de menor capacidad que atraviesa el flujo antes de entrar a la red OBS, el tiempo de ensamblaje de ráfagas (T_s) y el tamaño máximo de la ventana de TCP (W_m).

3.1. Flujos rápidos

Se define un flujo rápido como aquél en el que se cumple que $\lambda \cdot T_s \geq W_m$. Para tratar esta clase de flujos, se supondrá que todos los paquetes de una misma ventana de transmisión de TCP se ensamblan en una misma ráfaga a la entrada de la red OBS. Cada paquete puede ser descartado individualmente en dos situaciones:

a) En la parte eléctrica de la red, con probabilidad p_{dpe}.

b) En el nodo de ingreso a la red OBS, si el paquete llega en un intervalo de ráfaga en el que ya se ha alcanzado el número máximo de paquetes para una ráfaga, y por lo tanto no hay recursos en el nodo para memorizarlo; esto sucede con una probabilidad p_{dpo} dada por (2). Dado que se considera que todos los paquetes de una misma ventana de transmisión se ensamblan en la misma ráfaga, eso significa que si un paquete se descarta en el nodo de ingreso, todos los paquetes siguientes de la misma ventana serán descartados por la misma causa.

En muchos modelos de TCP se refleja la correlación en el descarte de paquetes mediante la aproximación pesimista (para una red EPS) de que, en una misma ventana, todos los paquetes siguientes a un paquete descartado son también descartados. En el caso del nodo de ingreso a la red OBS, según la hipótesis anterior, esta aproximación deja de ser pesimista, para reflejar el comportamiento establecido. En lo que sigue se aplicará esa aproximación tanto a la parte eléctrica como a la entrada de la parte óptica, que se considerarán conjuntamente, con una probabilidad de descarte de paquete $p_{\mathrm{dp}} = 1 - (1 - p_{\mathrm{dpe}})(1 - p_{\mathrm{dpo}})$ (suponiendo independencia entre sucesos).

Por otra parte, en el interior de la red OBS, una ráfaga se descarta con una probabilidad p_{dr}, y con ella se descartarán todos los paquetes pertenecientes a la misma ventana de transmisión de TCP.

Hechas las puntualizaciones anteriores, se desarrolla a continuación el modelo centrado en la versión Reno de TCP, que incluye mecanismos de comienzo lento, retransmisión rápida y recuperación rápida.

Se considera el flujo dividido en rondas; una ronda equivale a la transmisión de una ventana completa. Cada vez que se recibe un asentimiento acumulativo de b paquetes en una ventana de tamaño W paquetes, se incrementa en $1/W$ (incremento lineal de pendiente $1/b$ por cada *Round Trip Time* – RTT). El decrecimiento de la ventana de transmisión puede ser provocado por dos eventos: la recepción de un asentimiento duplicado (TD – *Triple Duplicate Acknowledgement*), en cuyo caso el tamaño disminuye a la mitad, y se continúa con la etapa de incremento lineal; y el vencimiento de una temporización (TO – *TimeOut*), en cuyo caso el tamaño pasa a ser de un paquete, y se hace *slow-start* hasta alcanzar la mitad del último tamaño de ventana (a la recepción del evento TO), para luego volver al incremento lineal.

El objetivo del modelo es hallar una fórmula para el cálculo de la tasa de transmisión del flujo en régimen permanente B. Dividiendo el flujo en períodos fundamentales —constituidos por una serie de rondas según se explica más adelante—, si en cada uno de ellos se transmiten Y_i paquetes durante un tiempo A_i, se tiene: $B = \frac{E[Y]}{E[A]}$.

3.1.1. Modelo de TDs

Se define un período fundamental LI (LIP) como el transcurrido durante la fase de incremento lineal entre dos indicaciones de pérdida. La secuencia $\{W_i\}$ representa los tamaños de ventana al final de cada LIP. Como aproximación, para poder simplificar el modelo analítico, se supondrá que el valor W_i es constante e igual a su media ($E[W]$). Un LIP comienza tras decrementar la ventana a la mitad ($W_{\mathrm{i\text{-}1}}/2$). Si α_i es el primer paquete perdido en el LIP$_i$, y X_i, la ronda en la que ocurre, habrá una última ronda con $W_i - 1$ paquetes antes de que ocurran un TD o un TO. Así:

$$E[Y_{\mathrm{LIP}}] = E[\alpha] + E[W] - 1. \qquad (5)$$

Llamando γ_i al número de paquetes que se envían en la última ronda de un LIP (igual al número de paquetes correctamente asentidos en la penúltima ronda) y teniendo en cuenta que $W_i = \frac{W_{\mathrm{i\text{-}1}}}{2} + \frac{X_i}{b}$ y que, por lo tanto, $E[W] = \frac{2}{b}E[X]$, se puede llegar a la ecuación (6).

$$E[\alpha] = E\left[\frac{X_i}{2}\left(W_i + \frac{W_{\mathrm{i\text{-}1}}}{2} - \frac{3}{b}\right) - \frac{W_{\mathrm{i\text{-}1}}}{2} + \frac{1}{b} + \gamma_i + 1\right] = \frac{E[X]}{2} \cdot \left(\frac{3E[W]}{2} - \frac{3}{b}\right) - \frac{E[W]}{2} + \frac{1}{b} + E[\gamma] + 1 \qquad (6)$$

El número de paquetes que se transmiten en la última ronda de un LIP (γ) posee una distribución geométrica condicionada al descarte de ráfaga; su media vendrá dada por la ecuación (7), en la que se ha supuesto $E[W]$ entero y se ha simplificado.

Sólo resta calcular la media de W, a partir de la de X. Dado un tamaño de ventana w en una ronda determinada, la probabilidad de que se produzca algún descarte en ella es:

$$p_d[w] = p_{\mathrm{dr}} + (1 - p_{\mathrm{dr}})[1 - (1 - p_{\mathrm{dp}})^w].$$

Así, se puede obtener fácilmente la distribución de X en función del tamaño de ventana ($E[W]$) al final del LIP anterior, y con ella se llega a la esperanza de W, ecuación (8), ecuación implícita que puede ser resuelta numéricamente para hallar $E[W]$. Conocido este valor, se puede calcular el número medio de paquetes transmitidos en un LIP, ecuación (9), sustituyendo (7) en (6), y ésta en (5).

En cuanto a la duración de un LIP, ya que $E[W] = \frac{2}{b}E[X]$, si se considera el RTT constante se puede estimar como:

$$E[A_{\mathrm{LIP}}] = \left(\frac{b \cdot E[W]}{2} + 1\right) RTT.$$

3.1.2. Modelo de TOs

Cuando se produce una pérdida y la última ronda del LIP no contiene tres o más paquetes, no se puede producir un TD, así que se retransmite tras una temporización (T_0), disminuyendo la ventana de congestión hasta un paquete. Antes de cada retransmisión, los tiempos de espera se incrementan de forma exponencial hasta un máximo de $64T_0$, siendo a partir de aquí el período de retransmisión constante. El intervalo de tiempo desde que se detecta el primer TO hasta que se reanuda la transmisión normal se denomina período TO (TOP).

El número de paquetes transmitidos durante un intervalo TO, si un paquete (que ahora coincide con la propia ventana de transmisión) se descarta con una probabilidad $p = 1 - (1 - p_{\mathrm{dr}})(1 - p_{\mathrm{dp}})$, sigue una distribución geométrica de media

$$E[Y_{\mathrm{TOP}}] = \frac{1}{(1 - p_{\mathrm{dr}})(1 - p_{\mathrm{dp}})}.$$

La duración media es de:

$$E[A_{\mathrm{TOP}}] = T_0 \frac{1 + p + 2p^2 + 4p^3 + 8p^4 + 16p^5 + 32p^6}{1 - p}.$$

3.1.3. Modelo Completo

Si se considera el régimen permanente como una sucesión de períodos fundamentales compuestos cada uno de ellos por n_i LIP y un TOP, de modo que n_i sea independiente de la duración de cada uno de estos subperíodos, se puede escribir:

$$B = \frac{E[Y_{\mathrm{LIP}}] + Q \cdot E[Y_{\mathrm{TOP}}]}{E[A_{\mathrm{LIP}}] + Q \cdot E[A_{\mathrm{TOP}}]}$$

donde Q representa la probabilidad de que la indicación de pérdida en la que termina un LIP sea un TO ($Q = 1/E[n]$). Para el cálculo de Q, se define $A(w,k)$ como la probabilidad de que los k primeros paquetes sean asentidos en una ronda de w paquetes, condicionada a que se haya producido alguna pérdida en dicha ronda:

$$A(w,k) = \begin{cases} \frac{p_{\mathrm{dr}} + p_{\mathrm{dp}}(1-p_{\mathrm{dr}})}{p_{\mathrm{dr}} + [1-(1-p_{\mathrm{dp}})^w](1-p_{\mathrm{dr}})}, & k = 0 \\ \frac{(1-p_{\mathrm{dp}})^k p_{\mathrm{dp}}(1-p_{\mathrm{dr}})}{p_{\mathrm{dr}} + [1-(1-p_{\mathrm{dp}})^w](1-p_{\mathrm{dr}})}, & 0 < k < w. \end{cases}$$

Si, además, se define $C(n,m)$ como la probabilidad de que m paquetes sean asentidos en secuencia en la última ronda (en la que se enviaron n paquetes) y el resto se pierdan:

$$C(n,m) = \begin{cases} p_{\mathrm{dp}}(1 - p_{\mathrm{dr}}) + p_{\mathrm{dr}}, & m = 0 \\ (1 - p_{\mathrm{dp}})^m p_{\mathrm{dp}}(1 - p_{\mathrm{dr}}), & 0 < m < n \\ (1 - p_{\mathrm{dp}})^n (1 - p_{\mathrm{dr}}), & m = n. \end{cases}$$

Entonces, la probabilidad de que una pérdida en una ventana de tamaño w sea un TO viene dada por la ecuación siguiente:

$$Q(w) = \begin{cases} 1, & w \leq 3 \\ \sum_{k=0}^{2} A(w,k) + \\ \sum_{k=3}^{w} A(w,k) \sum_{m=0}^{2} C(k,m), & w > 3. \end{cases}$$

Para evitar las discontinuidades que posee la ecuación anterior, se puede suponer w entero y simplificar, para llegar a la ecuación (10).

$$\begin{aligned} E[\gamma] &= \sum_{k=1}^{E[W]-1} k \cdot \frac{(1 - p_{\mathrm{dr}})(1 - p_{\mathrm{dp}})^k p_{\mathrm{dp}}}{p_{\mathrm{dr}} + [1 - (1 - p_{\mathrm{dp}})^{E[W]}](1 - p_{\mathrm{dr}})} \\ &= \frac{(1 - p_{\mathrm{dr}})(1 - p_{\mathrm{dp}} - (1 - p_{\mathrm{dp}})^{E[W]}(E[W] - (E[W] - 1)(1 - p_{\mathrm{dp}})))}{(p_{\mathrm{dr}} + (1 - (1 - p_{\mathrm{dp}})^{E[W]})(1 - p_{\mathrm{dr}}))p_{\mathrm{dp}}} \end{aligned} \tag{7}$$

$$E[W] = \frac{2}{b}E[X] = \frac{2}{b}\sum_{k=1}^{\infty} k \cdot p_d\left[\frac{E[W]}{2} + \frac{k}{b}\right] \cdot \prod_{j=0}^{k-1}\left(1 - p_d\left[\frac{E[W]}{2} + \frac{j}{b}\right]\right) \tag{8}$$

$$E[Y_{\mathrm{LIP}}] = \frac{b \cdot E[W]}{4} \cdot \left(\frac{3E[W]}{2} - \frac{1}{b}\right) + \frac{1}{b} + \frac{(1 - p_{\mathrm{dr}})(1 - p_{\mathrm{dp}} - (1 - p_{\mathrm{dp}})^{E[W]}(E[W] - (E[W] - 1)(1 - p_{\mathrm{dp}})))}{(p_{\mathrm{dr}} + (1 - (1 - p_{\mathrm{dp}})^{E[W]})(1 - p_{\mathrm{dr}}))p_{\mathrm{dp}}} \tag{9}$$

$$Q(w) = \frac{(p_{\mathrm{dr}} + p_{\mathrm{dp}}(3 - (3 - p_{\mathrm{dp}})p_{\mathrm{dp}})(1 - p_{\mathrm{dr}}))(2 - (1 - p_{\mathrm{dr}})((1 - p_{\mathrm{dp}})^{w+1} + p_{\mathrm{dp}}(3 - p_{\mathrm{dp}}(3 - p_{\mathrm{dp}}))) - p_{\mathrm{dr}})}{1 - (1 - p_{\mathrm{dp}})^w(1 - p_{\mathrm{dr}})} \tag{10}$$

Suponiendo W constante e igual a $E[W]$, dado por (8), se puede aproximar $Q \simeq Q(E[W])$.

3.1.4. Comentarios

No se ha tenido en cuenta en este modelo el período de comienzo lento, considerándolo de duración despreciable; tampoco se ha tenido en cuenta el impacto de la limitación del tamaño de ventana, aunque no sería difícil de incorporar al mismo. Se puede observar que la ecuación (8) se simplifica notablemente si la probabilidad de descarte de ráfaga domina sobre la de paquete ($p_{\mathrm{dr}} \gg p_{\mathrm{dp}}$). En ese caso, se puede obtener una ecuación explícita:

$$E[W] = \frac{2(1-p_{\mathrm{dr}})}{b \cdot p_{\mathrm{dr}}}.$$

Sin embargo, los resultados obtenidos en este caso son, en general, poco exactos. Ello es debido a que si domina el descarte de ráfagas, la mayor parte de las detecciones de pérdida serán TOs (Q elevado), con lo que la recuperación rápida no surtirá efecto, y el comportamiento de TCP Reno quedará reducido al de TCP Tahoe (no se ha tenido en cuenta la retransmisión rápida). Además, si el tamaño de ventana es relativamente elevado, la fase de comienzo lento será relativamente larga, y tendrá un peso considerable en el valor final de *throughput* alcanzado. Sin embargo, si el factor dominante es p_{dp}, para que Q sea elevado, es necesario que p_{dp} también lo sea; así, el tamaño de ventana será relativamente pequeño, y la fase de comienzo lento seguirá siendo despreciable. Esta es la razón fundamental que justificaba la omisión de la fase de comienzo lento en el modelo de [5], que deja de ser válida para pérdidas elevadas en el interior de la red OBS.

3.2. Flujos lentos

Son flujos lentos aquéllos para los que $\lambda \cdot T_s < 1$. En esta clase de flujos se supondrá que todos sus paquetes viajan a través de la red OBS en ráfagas distintas, de modo que los descartes se pueden considerar independientes, con una probabilidad $p = 1-(1-p_{\mathrm{dr}})(1-p_{\mathrm{dp}})$.

Al postularse esta hipótesis, el modelo de una red de paquetes es perfectamente aplicable a este caso, como por ejemplo el de [5], teniendo en cuenta la modificación en la probabilidad de descarte según la ecuación anterior. También se puede aplicar el modelo de flujos rápidos, corrigiendo los parámetros de entrada:

$$p'_{\mathrm{dr}} = 0, \quad p'_{\mathrm{dp}} = 1-(1-p_{\mathrm{dr}})(1-p_{\mathrm{dp}}).$$

3.3. Flujos de tasa media

Para estos flujos se cumple que $1 \leq \lambda \cdot T_s < W_m$. Su comportamiento es intermedio entre los de las clases anteriores, así que, en función de su proximidad con una u otra se podrían emplear los modelos anteriores. El tratamiento matemático de esta clase de flujos es más complejo que el de los rápidos, pero se puede conseguir una aproximación utilizando el mismo modelo de flujos rápidos, corrigiendo los parámetros de entrada según la siguiente fórmula, que representa una recta entre $p'_{\mathrm{dr}} = 0$ y $p'_{\mathrm{dr}} = p_{\mathrm{dr}}$, manteniendo constante la probabilidad global de descarte:

$$p'_{\mathrm{dr}} = p_{\mathrm{dr}} \frac{\lambda \cdot T_s - 1}{W_m - 2}$$
$$p'_{\mathrm{dp}} = 1 - \frac{(1-p_{\mathrm{dr}})(1-p_{\mathrm{dp}})}{1-p'_{\mathrm{dr}}}.$$

3.4. Comprobación experimental

Para la comprobación experimental se ha utilizado un modelo de pérdidas adaptado a las hipótesis consideradas, manteniendo fijo el *Round Trip Time* (RTT) en 800 ms, más el tiempo de ensamblado de ráfagas en los nodos de ingreso a la red OBS, que se ha establecido en 10 ms; el tamaño de paquete se ha fijado en 1000 bytes.

Para flujos rápidos se ha utilizado un enlace de entrada a la red OBS de $8 \cdot 10^{12}$ bps, que permite introducir en la misma $8 \cdot 10^{12} \cdot 10^{-2}/8000 = 10^7$ paquetes durante un intervalo de ensamblado de ráfaga; los resultados obtenidos se muestran en las Figs. 2 y 4, que representan, respectivamente, el *throughput* en función de p_{dp} manteniendo constante p_{dr}, y en función de p_{dr}, manteniendo constante p_{dp}. En las Figs. 3 y 5 se muestran los tamaños de ventana medios ($E[W]$) en las mismas condiciones anteriores.

En las gráficas se puede observar que el modelo es bastante preciso en los rangos de probabilidades de descarte moderadas-bajas (entre 0,001 y 0,1 los resultados son muy precisos).

En cuanto a los resultados obtenidos, se puede ver que el comportamiento del flujo ante la variación de p_{dr} es mucho más suave que ante la variación de p_{dp}, siendo además la disminución de la tasa de transmisión menos notable ante el incremento de p_{dr} que ante el de p_{dp}. Esto se debe en parte a la hipótesis de correlación de las pérdidas de paquetes, mientras que las pérdidas de ráfagas se han supuesto independientes. A pesar de todo, ante pérdidas elevadas de ráfagas, la tasa decrece rápidamente, pues la pérdida de una ráfaga supone una detección por TO, y la reducción de la ventana de congestión a un paquete.

Dada la corrección introducida en el modelo para flujos lentos, según la cual la probabilidad de descarte de ráfagas se traduce en pérdidas de paquetes, se puede predecir que para estos flujos con probabilidades de descarte moderadas-bajas, la tasa resultante será menor que para los rápidos.

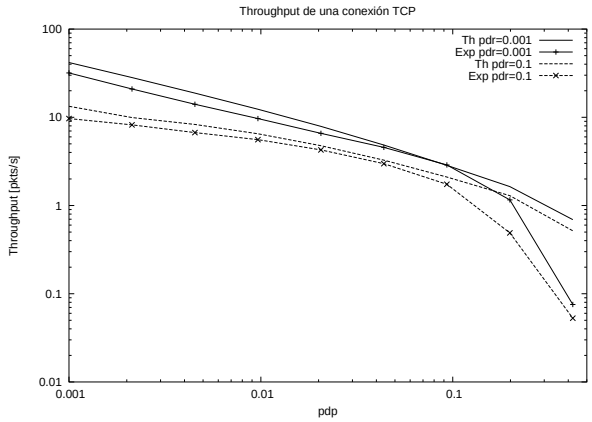


Figura 2: *Throughput* variando p_{dp} (Flujo rápido).

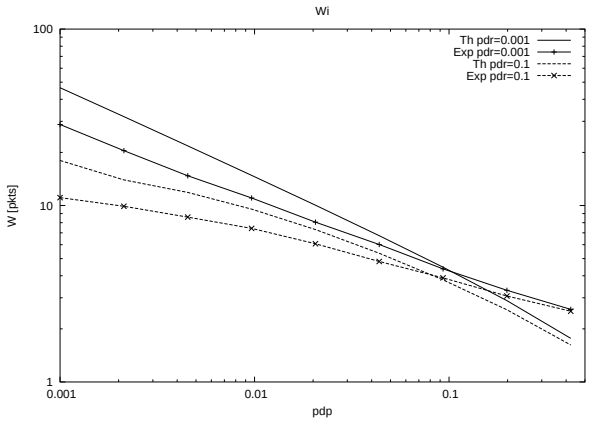


Figura 3: $E[W]$ variando p_{dp} (Flujo rápido).

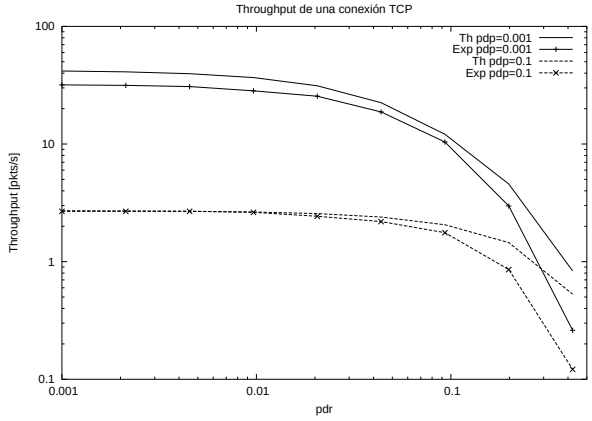


Figura 4: *Throughput* variando p_{dr} (Flujo rápido).

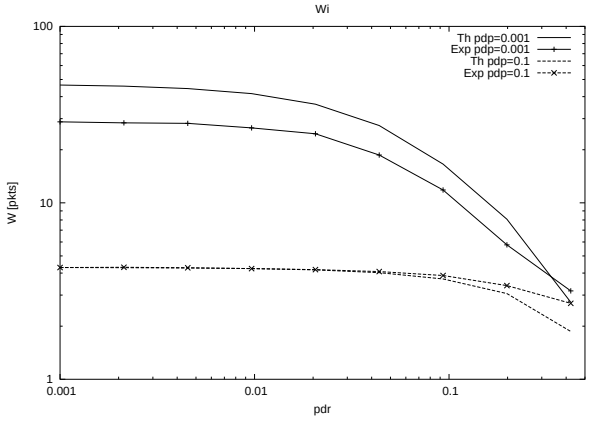


Figura 5: $E[W]$ variando p_{dr} (Flujo rápido).

Las pruebas se han realizado con un tiempo de ensamblado de ráfaga de 1ms y un enlace de ingreso a la red OBS de capacidad $8 \cdot 10^6$ bps, con lo que en cada ráfaga entrará un paquete de este flujo.

Los resultados obtenidos se muestran en las Figs. 6, 7, 8 y 9, que son análogas a las presentadas para flujos rápidos. En ellas se ha incluido además la predicción dada por el modelo de Padhye *et al.* [5]. Como se puede ver, efectivamente los flujos lentos presentan un comportamiento peor que los rápidos y, además, el modelo desarrollado en el presente texto proporciona una aproximación relativamente mejor que el modelo de [5].

4. Conclusiones

Se ha presentado un nuevo modelo de pérdidas aplicable a una red OBS, y un nuevo modelo de la tasa de envío de un flujo TCP que atraviesa una red OBS. Estas dos herramientas posibilitan realizar un análisis teórico completo de una red OBS. La validez de ambos modelos ha sido verificada mediante simulaciones con *ns-2*.

Del estudio de los resultados se desprende que las pérdidas de ráfagas en el interior de la red OBS son menos perjudiciales que las pérdidas de paquetes en una red EPS convencional, debido principalmente a que se han supuesto los descartes independientes del tamaño de ráfaga; así, mientras la probabilidad de que haya una reducción de ventana (de tamaño w) con descartes de ráfagas es de p_{dr}, con descartes de paquetes es de $1-(1-p_{\mathrm{dp}})^w$; por lo tanto, aunque ante un descarte de ráfaga se produce una reducción drástica (por TO) de la ventana de congestión de TCP, a igual proporción de descartes, el número de veces que la ventana se reduce es menor que en el caso de descartes de paquetes. Se podría pensar en un compromiso entre ambos efectos negativos, en flujos de tasa media; sin embargo, con las hipótesis realizadas, no se ha observado tal compensación, dándose el *throughput* máximo en los flujos rápidos.

Otro factor responsable, aunque en menor medida, del efecto anterior es la independencia entre pérdidas en el interior de la red OBS. A pesar de que estas redes también pueden sufrir congestión, por lo general sus ciclos de comportamiento son más dinámicos que los de una red EPS, debido fundamentalmente a la ausencia de *buffers*, por lo

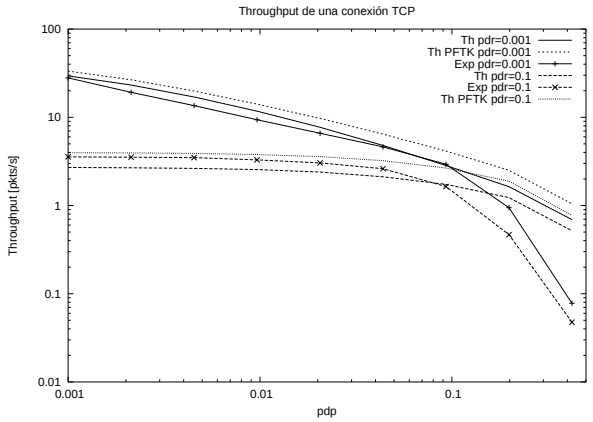


Figura 6: *Throughput* variando p_{dp} (Flujo lento).

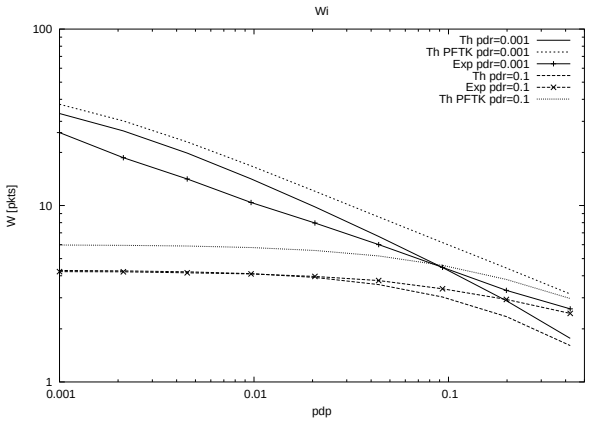


Figura 7: $E[W]$ variando p_{dp} (Flujo lento).

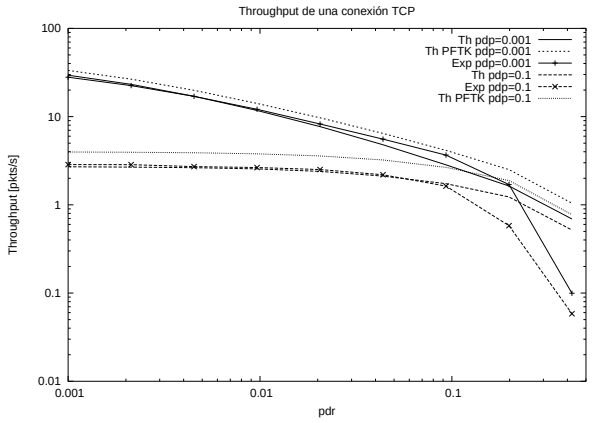


Figura 8: *Throughput* variando p_{dr} (Flujo lento).

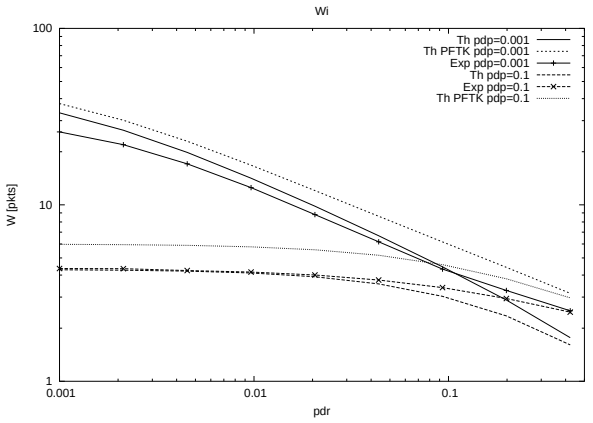


Figura 9: $E[W]$ variando p_{dr} (Flujo lento).

que la recuperación es más rápida. Además, como se puede observar de los resultados del modelo de pérdidas, el ensamblado de los paquetes entrantes en ráfagas suaviza el tráfico ofrecido a la red, y redunda en un comportamiento de ésta mejor del esperado en el caso de tráfico poissoniano. A la vista de los resultados, una variación en el mecanismo de control de congestión de TCP para adaptarlo a las peculiaridades de las redes OBS puede desembocar en un aprovechamiento mucho mayor de la capacidad de transmisión.

Como trabajo futuro de modelado, resta incorporar el efecto de la limitación del tamaño de ventana, así como el cálculo del *goodput*, o tasa real de transmisión ofrecida a las aplicaciones. También es necesario modelar la fase de comienzo lento para tener resultados más precisos en el caso de que los descartes sean únicamente de ráfagas.

Referencias

[1] A. A. Abouzeid, S. Roy, M. Azizoglu. "Stochastic modeling of TCP over lossy links". IEEE INFOCOM 2000 - The Conference on Computer Communications, no. 1, March 2000 pp. 1724-1733

[2] Y. Chen, C. Qiao and X. Yu. "Optical Burst Switching: A New Area in Optical Networking Research". *IEEE Network* May/June 2004 pp. 16-23.

[3] M. Izal, J. Aracil. "On the influence of self similarity on Optical Burst Switching traffic". Proc. IEEE GLOBECOM 2002.

[4] L. Kleinrock. *Queueing Systems. Volume I: Theory.* John Wiley & Sons, 1975

[5] J. Padhye, V. Firoiu, D. F. Towsley, J. F. Kurose. "Modeling TCP reno performance: A simple model and its empirical validation". *IEEE/ACM Transactions on Networking*, no. 2, Apr 2000 pp. 133-145.

[6] B. Sikdar, S. Kalyanaraman, K. S. Vastola. "Analytic models for the latency and steady-state throughput of TCP Tahoe, Reno, and SACK". *IEEE/ACM Transactions on Networking*, no. 6, Dec 2003 pp. 959-971

[7] W. R. Stevens. *TCP/IP Illustrated. Vol I: The Protocols.* Addison Wesley, 1994.

[8] X. Yu, C. Qiao and Y. Liu. "TCP Implementations and False Time Out Detection in OBS Networks". IEEE INFOCOM 2004.

Diseño y evaluación de un protocolo de descubrimiento de servicios para redes móviles Ad hoc

Mª Isabel Vara Lorenzo[1], José Mª Cabero López[1],
José Luis Jodrá Luque[2], José Oscar Fajardo Portillo[2]
[1]Unidad de TELECOM. Fundación ROBOTIKER
[2]Departamento de Ingeniería Telemática. Universidad del País Vasco

[1]{mvara,jmcabero}@robotiker.es
[2]jtpjoluj@bi.ehu.es, jbtfapoj@bipt106.bi.ehu.es

Abstract*. This paper proposes a service discovery protocol for discovering and advertising services in a proactive ad hoc network. The protocol we have defined is piggybacked into the OLSR (Optimized Link State Routing) routing protocol. We have defined a new message type into OLSR, called Service Discovery Message (SDM), for both advertisement and discovery of services. The advertisement frequency and advertisement lifetime are user-controlled parameters, so that they can be modified depending on the user requirements. Each node maintains a service cache to store information about its own services, and the services each device discovers in the network. We also present simulation results of our protocol and show that the service discovery protocol defined here, achieves much efficiency in discovering services, while it introduces practically no packet overhead compared to the basis OLSR protocol.*

1 Introducción

Una red móvil ad hoc (Mobile Ad hoc Network o MANET) se define como una red de comunicaciones, que se forma cuando se necesita y que no precisa de ningún tipo de infraestructura de red fija. Esto permite que se cree una red prácticamente de la nada, sin necesidad de intervención humana ni configuraciones previas.

Los nodos que conforman la red participan en la toma de decisiones y tienen su propio conjunto de protocolos de enrutamiento, en el que toman parte de forma activa. El tiempo de vida de estos nodos es generalmente corto y, entran y salen de la red sin previo aviso. Esto hace que la topología de la red cambie de forma dinámica y aleatoria.

La mayoría de los nodos son dispositivos con limitado poder de procesamiento y baja capacidad de almacenamiento de energía. Estos dispositivos disponen de servicios muy específicos como, por ejemplo, medir la temperatura de una sala, regular la iluminación y/o cambiar la música ambiental que, ofrecen al resto de usuarios de la red para que hagan uso de ellos. En la figura 1 se describe una posible red ad hoc.

En este entorno tan dinámico formado por redes ad hoc, donde multitud de equipos ofrecen multitud de servicios, un mecanismo de descubrimiento de servicios, facilitaría a un usuario acceder a los servicios que otro dispositivo ofrece en la red.

El principal objetivo del descubrimiento de servicios es, evitar o al menos minimizar miles de configuraciones cuando un usuario quiere conectarse a una red, buscar un equipo o acceder a prestaciones que éste ofrece. Sin embargo, en un entorno tan dinámico, en el que no existe una infraestructura central, diseñar un protocolo de descubrimiento de servicios eficiente, no resulta tarea fácil. Lo ideal sería implementar mecanismos que permitieran tener un conocimiento total de los servicios que se ofrecen en la red, y saber cuándo dejan de estar disponibles, sin que ello implique saturar de paquetes la red. De esta forma el ancho de banda que se consume en la red permanece prácticamente inalterado y se consigue también, alargar el consumo de las baterías de los dispositivos.

Nuestro trabajo actual se basa en desarrollar un protocolo de descubrimiento de servicios integrado en el protocolo de enrutamiento OLSR [1] que, soporte tanto el descubrimiento como el anuncio de servicios. Para ello, tomando como punto de partida el formato de los paquetes del protocolo OLSR, hemos descrito un nuevo tipo de mensaje denominado *Service Discovery Message (SDM)* que permite descubrir nuevos servicios a la vez que el protocolo OLSR descubre nuevas rutas. De esta forma aprovechamos la capacidad que tiene el protocolo OLSR para descubrir rutas, para obtener información de los servicios que ofrece cada nodo.

En la sección 2 realizaremos un breve estado del arte sobre los protocolos existentes actualmente para el descubrimiento de servicios en redes ad hoc. Después en la sección 3 realizaremos una breve revisión del protocolo OLSR. En la sección 4 describimos la propuesta para el descubrimiento de servicios sobre el protocolo OLSR que hemos realizado. En la sección 5 presentamos los resultamos que hemos obtenido. Por último, finalizamos con las conclusiones y las líneas futuras de nuestro trabajo, sección 6.

Fig. 1 Arquitectura de una red ad hoc

2 Estado del arte

2.1 Descubrimiento de servicios

El descubrimiento de servicios en redes ad hoc ha tenido un especial impacto en los últimos años. Pensar que un dispositivo cuando se incorpora a una red es capaz de localizar los recursos disponibles en ella resulta atractivo. Sin embargo, a pesar de su importancia, las soluciones que se han aportado han tenido pocos avances, y todos los resultados que se han obtenido se basan en simulaciones en entornos con pocos nodos.

Mecanismos de descubrimiento de servicios afianzados ya en el mercado como *SLP* [4], *Jini* [5], *UpNP* [6], *Salutation* y *Salutation Lite* [7], [8] y el *SDP de Bluetooth* [9], no son muy adecuados para las redes ad hoc. Aunque en todos ellos se ha tenido en cuenta el dinamismo del entorno, la mayoría no ha considerado las limitaciones de los dispositivos que conforman la red, ni la posibilidad de que las redes ad hoc que se forman entre ellos de forma espontánea, no incluyan ningún elemento fijo sin limitaciones tipo PC. Así, en la mayoría de ellos existe un servidor central en el que los nodos que ofrecen los servicios se registran y al que los clientes preguntan cuando necesitan un determinado servicio.

También han surgido otros protocolos y mecanismos especialmente diseñados para el descubrimiento de servicios, teniendo en cuenta las limitaciones de las redes ad hoc. El objetivo principal que todos se plantean es facilitar a un dispositivo móvil el descubrimiento, configuración y uso de los servicios que se ofrecen en la nueva red a la que se conectan. Entre todos destacan, Konark [10], que más que un protocolo de descubrimiento de servicios en sí, lo que se propone es un framework para el descubrimiento de servicios. Define también, cómo debe ser el acceso a los servicios; GSD [11], cuya principal novedad es que define los servicios en una ontología definida en DAML *(DARPA Agent Markup Language)* [12], de tal forma que las búsquedas de servicios se realizan según correspondencias semánticas y no sintácticas; DEAPspace [13], que es un protocolo de descubrimiento de servicios para operar en redes ad hoc de un solo salto. Los dispositivos tienen una *"world view"* del entorno que los rodea y la transmiten cada cierto tiempo a sus vecinos mediante mensajes broadcast; y por último Allia [14], que es un mecanismo de descubrimiento de servicios basado en agentes

En general, estas propuestas abordan el problema del descubrimiento de diferentes formas, desde las que han sido diseñadas para su uso con un determinado protocolo de red, hasta las que están asociadas a un determinado lenguaje de programación.

2.2 Integrar mecanismos de descubrimiento y anuncio de servicios con protocolos de rutado

La mayoría de las propuestas que existen para descubrimiento de servicios en una red ad hoc, separa el mecanismo de rutado del mecanismo de descubrimiento. El rutado se sitúa dentro de la capa de red, y el descubrimiento es una utilidad que se usa por encima de esta capa. A pesar de que esta solución resulta útil en una red cableada, en una red tan dinámica como es la red ad hoc, se ha comprobado que integrar el descubrimiento de servicios dentro de la capa de rutado aumenta la eficiencia de la red [15], [16].

En la actualidad existen varias propuestas que tienen en cuenta esto: GSR (Group-based Service Routing Protocol) [17] que hace uso de la infraestructura de GSD para transmitir los mensajes, en lugar de calcular rutas nuevas con un protocolo de rutado tradicional, [18], que añade extensiones al protocolo ODMRP [19] para permitir el descubrimiento de servicios. El protocolo AODV [20], [21] también ha definido una extensión del protocolo, que incluye el descubrimiento de servicios. En OLSR ha aparecido recientemente una propuesta para descubrimiento de servicios especialmente diseñada para aplicaciones SIP [22].

3 Descripción del protocolo OLSR

OLSR tiene como principales características el ser un protocolo de enrutamiento proactivo [2] basado en estado de enlace. Esto implica que los nodos que forman parte de la red ad hoc, intercambian periódicamente mensajes de control, que permiten aprender la topología de la red. El protocolo tiene la ventaja de reaccionar rápidamente ante cambios en la topología de la red, y por eso los nodos pueden elegir la mejor ruta en cada instante, dependiendo siempre de la frecuencia con la que se intercambien información. Cuando más frecuente sea el intercambio de mensajes de control, más actualizadas estarán las rutas, aunque esto implica sobrecargar mucho la red con mensajes de control. Y en redes como las redes ad hoc, donde el ancho de banda es limitado, inundar de mensajes la red no resulta eficiente.

Por lo tanto, al tratarse de un protocolo de enrutamiento por estado de enlace, los mensajes se retransmiten a toda la red. Esto resulta una operación costosa si tenemos en cuenta los limitados recursos de las redes móviles ad hoc.

Para reducir el número de retransmisiones redundantes que tienen lugar en una red ad hoc, el protocolo OLSR hace uso del mecanismo MPR [3]. Este mecanismo garantiza que los mensajes sigan llegando a todos los nodos, disminuyendo el número de nodos de la red que van a retransmitir los mensajes. El algoritmo que se sigue es el siguiente. Cada nodo calcula cuál es el número de vecinos mínimo que debería retransmitir su mensaje. Este conjunto mínimo de nodos es el que se conoce como *Multi Point Relays* (MPR). Este número se calcula, averiguando el menor número de vecinos que se necesitan, para alcanzar a todos los nodos que se encuentran a dos saltos de distancia. En la figura 2 se muestra cómo con el mecanismo MPR se reduce considerablemente el consumo de ancho de banda.

Para transportar los mensajes, OLSR define un formato de paquete básico que entienden todos los nodos que implementan el protocolo. Un ejemplo de este paquete se muestra en la figura 3. Se puede observar que el paquete puede contener varios mensajes diferentes. Cada mensaje comparte una cabecera común donde se indica el tipo de mensaje que se encapsula. En la especificación de OLSR se definen tres tipos de mensajes. Nosotros vamos a definir uno nuevo, que haga posible el descubrimiento de servicios en una red ad hoc.

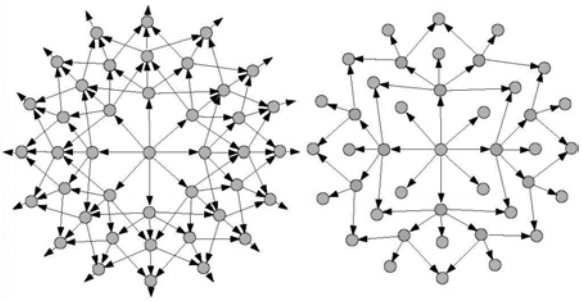


Fig. 2 Mecanismo de broadcast tradicional vs mecanismo MPR

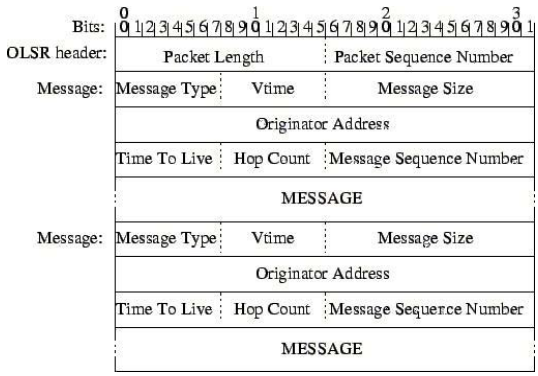


Fig. 3 Formato del paquete OLSR

4 Mecanismo de descubrimiento de servicios propuesto

El formato de paquete del protocolo OLSR, nos permite fácilmente añadir nuevas funcionalidades al protocolo, sin más que definir nuevos tipos de mensajes que irán encapsulados en el campo *MESSAGE*. Para permitir el descubrimiento de servicios, nosotros hemos definido un nuevo tipo de mensaje, *Service Discovery message (SDM)*, que describimos en la figura 4. Este mensaje permite a un dispositivo tanto descubrir servicios que ofrecen otros nodos, como anunciar los servicios que un nodo posee.

A continuación detallamos el significado de los campos más significativos del mensaje. *TimeToLive* define el número de nodos por los que el mensaje puede viajar. Es decir, el mensaje no se retransmite indefinidamente, sino únicamente hasta que se llegue al valor del *TimeToLive*. Este parámetro lo especifica el dispositivo que anuncia sus servicios. Se trata de un mensaje, por lo tanto, que no viaja por toda la red, sino que tiene limitado el número de saltos al valor que se indique en este campo. De esta forma conseguimos reducir el consumo de ancho de banda. El campo *HopCount* contiene el número de nodos por los que el mensaje ha pasado. Se inicia a 0 y se va incrementando en uno a medida que va pasando por otro nodo. De esta forma se mantiene un contador que controla, el número de nodos que el mensaje ha atravesado. El valor máximo lo limita el campo *TimeToLive.*

El campo *Type* indica si el mensaje es una petición de servicios o un anuncio de servicios. *Service Name* nos dice el nombre del servicio que se anuncia o que se solicita. *SDM_INTERVAL* indica cada cuánto se envía un mensaje de anuncio de servicio. Se trata de un parámetro configurable por el usuario, de tal forma que se pueda variar dependiendo del entorno en el que se opere. Cuanto menor sea este parámetro, el mensaje se transmitirá más a menudo, y la sobrecarga que se introduzca en la red será mayor. El campo *LifeTime* indica durante cuánto tiempo el servicio estará disponible para el resto de nodos. Hay que tener en cuenta que los servicios pueden tener un tiempo de vida transcurrido el cual dejan de estar disponibles. Por último, el parámetro *Service Description* contiene información de los servicios que un nodo solicita o de los servicios que un nodo anuncia.

Fig. 4 Formato de paquete SDM

4.1 Caché de servicios

Los dispositivos almacenan sus servicios locales y los que van descubriendo, en una caché de servicios. En esta caché de servicios, existe un espacio para almacenar los servicios propios del nodo (servicios locales) y otro, para almacenar los servicios que van descubriendo.

Cuando un nodo recibe un mensaje de anuncio de servicios extrae la información necesaria del paquete que recibe, y almacena su contenido en el espacio habilitado para ello dentro de la caché de servicios.

Cada entrada en la caché de servicios, bien sea para servicios locales como para servicios que un nodo va descubriendo, sigue el siguiente formato:

<Provider-Address, Service-Description, LifeTime, Number of services, Number of updated services>

donde *Provider-Address* identifica el nodo al que pertenece el servicio, *Service-Description* contiene las principales características del servicio, *LifeTime* indica los segundos que se espera que el servicio esté disponible. Antes de que expire este temporizador el nodo proveedor del servicio debería restablecer este tiempo, si desea que el servicio siga estando disponible para el resto de usuarios. Por último, los campos *Number of Services* y *Number of updated Services* describen el número de servicios que se guardan en la caché de servicios. Sólo el último servicio almacenado en la caché tiene actualizados estos campos. El resto de servicios hacen caso omiso de ellos. Cuando se guarda un nuevo servicio los valores de *Number of services y Number of updated services* correspondientes al servicio anterior dejan de tener validez. Ambos campos se inician a 0 y cada vez que se guarda un nuevo servicio, el campo *Number of updated Services* se incrementa en 1 respecto al valor de la entrada anterior en la caché. Cuando vence el temporizador *SDM_INTERVAL* el contenido de este campo se copia en el campo *Number of Services*. De esta forma un nodo es capaz de saber cuándo se ha guardado un nuevo servicio en su caché. Esto es importante cuando un dispositivo anuncia sus servicios, puesto que únicamente se anuncian servicios nuevos. Si el contenido de ambos campos es diferente, significa que el nodo tiene uno o más servicios que anunciar. Cuando los ha anunciado, y vence el temporizador *SDM_INTERVAL* los dos campos se vuelven a iniciar al mismo valor.

La caché de servicios irá aumentando con el número de servios que se vayan guardando, hasta llegar a un punto en que se llene. Cuando ocurre esto, hay que ir borrando servicios para dar cabida a otros nuevos. Los primeros que se van a borrar son aquellos cuyo parámetro *LifeTime* esté a punto de expirar. De la misma forma, todos los servicios cuyo *LifeTime* haya expirado, se borran automáticamente de la caché de servicios.

4.2 Anuncio de servicios

Se trata de un mensaje que transmiten aquellos nodos que tienen algún servicio que anunciar. El formato del paquete es el que se describe en la figura 4.

Cada *SDM_INTERVAL,* fijo para todos los nodos que componen la red, cada nodo enviará un nuevo mensaje con los servicios nuevos que disponga. Únicamente se enviarán actualizaciones de los servicios o servicios nuevos. De esta forma se minimiza la sobrecarga que se introduce en la red. En caso de que el nodo no disponga de servicios nuevos, no se incluirá ningún mensaje *SDM* dentro de la cabecera del paquete OLSR. Este mecanismo hace que si se incorporan nuevos nodos a la red éstos no reciban los mensajes de anuncio de servicio de servicios que ya se han anunciado, y por lo tanto, tienen dos opciones: esperar a que el temporizador *LifeTime* venza, y el servicio se vuelva a anunciar, o bien, solicitar ellos mismos el servicio que deseen si nadie se lo ha anunciado.

La forma que tiene un nodo de saber si tiene o no nuevas versiones de servicios o servicios nuevos, es comprobando los valores de los campos *Number of services* y *Number of updated services* de su caché de servicios. A continuación se muestra un pseudo-código donde se realiza este proceso:

```
if ( SDM_INTERVAL )
    if(compare(Number of updated Services,
    Number of Services ) is different )
        then Advertise new services;
    else
        No services are transmitted
```

Cualquier dispositivo móvil que escuche un mensaje de anuncio de servicios y que esté interesado en los servicios que se anuncian guardará la información necesaria de los mismos en su caché de servicios.

Es posible que el mismo servicio lo oferten varios nodos. Como la caché de servicios de un dispositivo tiene un tamaño limitado, antes de guardar la información referente a un servicio, cada nodo comprueba si esa información ya ha sido previamente almacenada en su caché, porque haya sido ofrecida por otro dispositivo, nunca por el mismo. Recordemos que nunca un mismo dispositivo anunciará el mismo servicio, a no ser que sea una actualización del mismo (por ejemplo, cuando el *LifeTime* del servicio esté a punto de expirar), puesto que únicamente se anuncian nuevos servicios. En caso de que ese servicio ya haya sido ofertado por otro dispositivo y se guarde copia de la información en la caché, el nodo hará caso omiso a la información que se anuncia. El algoritmo que se sigue es el siguiente:

```
if (Duplicate(service))
    then Not store the service;
else
```

Extract from the advertisement the fields needed to store in the service cache.

4.3 Descubrimiento de servicios

El formato de un mensaje de descubrimiento de servicios es el mismo que el del anuncio de servicios y por tanto, el descrito en la figura 4. Lo único que cambia es que ahora el tipo de mensaje es de descubrimiento y el *ServiceName* identifica al servicio que queremos buscar. Los campos *SDM_INTERVA*L y *LifeTime* viajan vacíos, puesto que ahora no se trata de anunciar un servicio cada cierto tiempo, sino de descubrir un servicio y esto se hará cuando el dispositivo lo necesite.

Los nodos que son MPR son los encargados de retransmitir el mensaje. Cuando un nodo recibe este mensaje, comprueba si alguno de los servicios que él soporta se corresponde con el servicio que se busca. En caso afirmativo, prepara un mensaje de respuesta al nodo que hace la petición del servicio. Pero antes de enviar la respuesta, espera un tiempo aleatorio en el que comprueba, si algún otro nodo ha respondido al nodo origen porque también tuviera el servicio que se solicitaba. De ser así, el nodo ignora el mensaje y no lo responde. De esta forma se evitan múltiples respuestas ofreciendo el mismo servicio, con lo que se gana ancho de banda.

En caso de que no haya ningún nodo que haya respondido a ese mensaje, se rellenará el mensaje SDM y se enviará al destinatario. Este mensaje lo escuchan todos los nodos de la red. Esta es una característica de los protocolos proactivos. A pesar de que este mecanismo pueda sobrecargar la red, algo que en el caso del protocolo OLSR se mejora con el mecanismo MPR, tiene la ventaja de que como el resto de nodos también escuchan la respuesta, en caso de que estén interesados en el servicio, pueden almacenar en su caché una entrada con las características del mismo, y así evitar buscarlo, si en un futuro necesitan un servicio de este estilo. De hecho, cuanto más se conozcan los nodos entre sí, menos peticiones de servicio se harán y, en consecuencia, menos se sobrecargará la red con nuevas peticiones.

5 Análisis de resultados

5.1 Entorno de simulación

La evaluación del protocolo de descubrimiento de servicios que hemos implementado sobre el protocolo proactivo OLSR, la hemos realizado utilizando el simulador de redes ns-2 [[23]. Partimos de una implementación de OLSR existente en ns-2 [24] a la que añadimos el nuevo tipo de mensaje con las características descritas en el apartado 4. Los resultados que se presentan son únicamente preliminares, y nos sirven para saber si la opción que proponemos es viable o no. Nuestro objetivo es mejorar el protocolo propuesto para que se pueda utilizar como un mecanismo estándar de descubrimiento de servicios para las redes ad hoc.

En las simulaciones los nodos móviles se desplazan utilizando el modelo random waypoint [25]. El funcionamiento es el siguiente: inicialmente los nodos están estáticos durante un tiempo de pausa. Seleccionan un destino aleatorio dentro del área de simulación y se dirigen a él con una velocidad media de 10m/s. Una vez que alcanzan el destino repiten el proceso hasta el final de la simulación.

El canal radio para cada nodo móvil tiene una capacidad de 2Mbps y usa IEEE 802.11b como tecnología de transmisión inalámbrica. El rango de comunicación es de 250 metros. Las simulaciones se realizan sobre un escenario de 1000x1000metros, durante un tiempo de 900 segundos. Se simulan 50 nodos. Cada simulación se repite 10 veces y se toma la media aritmética del resultado. Variamos la frecuencia de envío del mensaje SDM para el anuncio de servicios entre 0,5s, 1,5s y 2,5s.

5.2 Resultados de la simulación

El principal objetivo es comprobar la viabilidad del protocolo propuesto y analizar si se introduce mucha sobrecarga en la red. También vamos a comparar el descubrimiento de servicios del protocolo OLSR con el que ofrece el protocolo ODMRP. Lo haremos en términos de tiempo que se tarda en descubrir un nuevo servicio.

En la figura 5 se muestra la sobrecarga que se introduce debida al nuevo mensaje SDM que hemos introducido. Se puede observar que la sobrecarga que se introduce cuando los mensajes SDM se envían cada 2,5s, es prácticamente la misma que para el protocolo OLSR básico, sin soporte para descubrimiento de servicios. Incluso cuando los mensajes SDM se envían con la máxima frecuencia, 0,5s, el tráfico que se genera es adecuado para una red MANET. Por lo tanto, podemos concluir que el exceso de tráfico que se genera al introducir este nuevo mensaje, es insignificante y por tanto, el protocolo se podría utilizar como primera aproximación, y a falta de pruebas más exhaustivas, como un mecanismo de anuncio y descubrimiento de servicios dentro del protocolo OLSR.

Otra de las pruebas que hemos realizado es comprobar el número de peticiones de descubrimiento de servicio que se hacen y el tiempo que se tarda en obtener una respuesta. Comparamos este tiempo con el que ofrece el descubrimiento de servicios del protocolo ODMRP. La figura 6 muestra los resultados que hemos obtenido. Se observa que el mecanismo de descubrimiento de servicio en el protocolo OLSR es mucho más eficiente que el de ODMRP. La principal razón es que en el protocolo OLSR tenemos un mayor conocimiento de la red y las rutas para transmitir los mensajes se conocen de antemano.

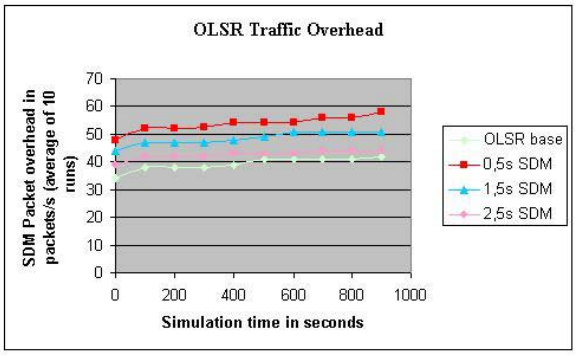


Fig. 5 Sobrecarga introducida en la red medida en paquetes/s para diferentes frecuencias de envío del mensaje SDM

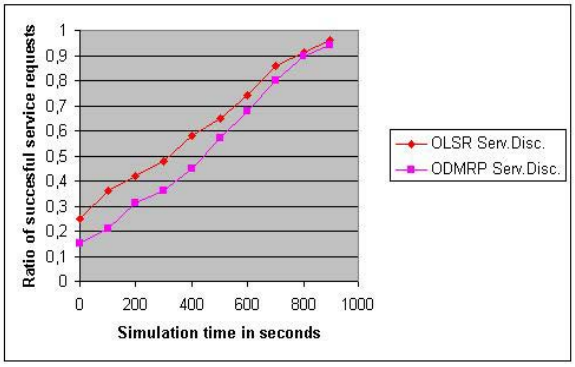


Fig. 6 Porcentaje de peticiones satisfactorias

6 Conclusiones y trabajo futuro

Hemos presentado un protocolo para descubrimiento de servicios en redes ad hoc proactivas basado en el protocolo de rutado OLSR.

Los resultados que hemos obtenido en las simulaciones demuestran que el protocolo descrito es viable para este tipo de redes. No introduce mucha sobrecarga y el proceso de descubrimiento de un nuevo servicio es relativamente corto.

Actualmente estamos comprobando cómo se comporta el protocolo cuando la caché de servicios se llena. También estamos verificando si el protocolo detecta cuándo un servicio deja de estar disponible, y cuánto tiempo tarda en averiguarlo.

Referencias

[1] T. Clausen and P.Jacquet "Optimized Link State Routing Protocol(OLSR". *RFC 3626*, IETF Network Working Group, October 2003.

[2] http://www.ietf.org/html.charters/manet-charter.html, work in progress.

[3] Amir Qayyum, Laurent Viennot, and Anis Laouiti. Multipoint Relaying for Flooding Broadcast Messages in Mobile Wireless Networks. *In Proceedings of the 35thHawaii International Conference on System Sciences*, 2002.

[4] E. Guttman, C. Perkins, J. Veizades, and M. Day, "Service Location Protocol, Version 2", RFC 2608, June 1999.

[5] K. Edwards and T. Rodden. Jini Example by Example. Prentice Hall PTR, June 2001.

[6] R.John. UpnP, Jini and Salutation- A Look at some popular Coordination Frameworks for Future Network Devices. Technical report, California Software Labs, 1999.

[7] Salutation Consortium. White Paper. Salutation Arquitecture Specification. http://www.salutation.org/whitepaper/originalwp.pdf,1998.

[8] Salutation Lite Homepage. http://www.salutation.org/lite/litesource.htm.

[9] Bluetooth Specification. Service Discovery Protocol (SDP). http://www.bluetooth.com, November 1999.

[10] S. Helal, N. Desai, V.Verma, C.Lee, "Konark: A Service Discovery and Delivery Protocol for Ad-Hoc Networks", WCNC, 2003.

[11] Anupam Joshi Dipanjan Chakraborty. GSD: A novel group-based service discovery protocol for MANETS. In *IEEE Conference on Mobile and Wireless Communications Networks, Stockholm, Sweden.*, September 2002.

[12] DARPA Agent Markup Language and Ontology Inference Layer. http://www.daml.org/2001/03/daml+oil.daml.

[13] R. Hermann, D. Husemann, M.Moser, M, Nidd, C. Rohner, A. Schade, "DEAPspace: Transient ad hoc networking of pervasive devices", Computer Networks, 2001.

[14] Olga Vladi Ratsimor *et al.*, "Allia: Alliance-based Service Discovery for Ad-Hoc Environments", InProceedings, *ACM Mobile Commerce Workshop*, September 2002.

[15] P. Bhagwat B. Raman and S. Seshan. Arguments for Cross-Layer Optimizations in Bluetooth Scatternets. *In the 2001Symposium on Applications and the Internet (SAINT)*, January 2001.

[16] E. Schwartz W. Adjiw-Winoto and H Balakrishnan. The Design and Implementation of an Intentional Naming System. *In Proceedings of the Symposium on Operating Systems Principles*, South Carolina, USA, December 1999.

[17] Dipanjan Chakraborty *et al.*, "Integrating Service Discovery with Routing and Session Management for Ad hoc Networks", *Ad Hoc Networks Journal by Elsevier Science*, March 2004.

[18] L. Cheng. "Service Advertisement and Discovery in Mobile Ad-Hoc Networks". In proceedings of Workshop on Ad hoc Communications and Collaboration in Ubiquitous Computing Environments (ACMCSCW 2002), ACM, 2002.

[19] Lee, S., Gerla, M., and Toh, C. On-demand multicast routing protocol (ODMRP) for ad hoc networks. Internet Draft, work in progress, June 1999.

[20] C.E. Perkings and E.M Royer. Ad-hoc On-Demand Distance Vector Routing. In 2nd IEEE Workshop on Mobile Computing Systems and Applications, pages 90-100, February 1999.

[21] C.E. Perkings and E.M Royer and S. Das. Ad hoc on-demand distance vector protocol. In IETF Internet Draft. Version 12,November 2002.

[22] L. Li, L. Lamont, "Service Discovery for Support of Real-time Multimedia SIP Applications Over OLSR MANETs", Li Li, L. Lamont, OLSR Interop & Workshop 2004, San Diego, USA, August 6-7, 2004.

[23] *The Network Simulator ns-2 Homepage*. http://www.isi.edu/nsnam/ns.

[24] *OLSR implementation for ns*-2 tool simulator *Homepage. http://ants.dif.um.es/masimum/um-olsr/html.*

[25] David B. Johnson and David A. Maltz, Dynamic Source Routing in ad hoc Wireless Networks, *in Mobile Computing*, edited by Tomas Timielinski and Hank Korth, Chapter 5, pages 153-181, Kluwer Academic Publishers, ISBN:0792396979, 1996.

Infraestructura para Servicios e Interfaces sensibles a la localización en Hogares Inteligentes

Miriam Machuca, Miguel A. López, Juan R. Velasco e Ivan Marsá Maestre
{miriam|miguellop|juanra|ivmarsa}@aut.uah.es
Departamento de Automática
Universidad de Alcalá
Edificio Politécnico Crtra. N-II Km. 31,600 28871 Alcalá de Henares.

Abstract *Some not widespread environments like sophisticated vehicles adjust controlled elements, like the seat and the rearview, in order to match the preferences of their users. In this context computer systems are fully capable of providing customized interfaces for users. However, this kind of service customization has not yet reached the home environment. Inside real home environments, we can find new services based on automatizing traditional ones, which make our lives easier and more comfortable. However, these services are provided independently, the degree of personalization is still very low, and the results are insufficient. The smart home must release the user from performing routine and tedious tasks to achieve comfort, security, and effective energy management. To achieve this goal, designed systems must use all posible components at home, providing a high quality service. In this paper we extend our previous work on using multiagent systems to build a smart home environment. We describe system functionality and introduce a new ontology in order to support communication and knowledge sharing.*

1 Introducción

Los entornos con los que interactuamos habitualmente, nuestro hogar, nuestro coche, nuestra oficina, tienden a ofrecernos un nivel de confort en continuo crecimiento. Hemos introducido en nuestro hogar un número cada vez mayor de dispositivos electrónicos, que ofrecen nuevos servicios o automatizan los servicios tradicionalmente conocidos. Por ejemplo, empiezan a aparecer dispositivos que proporcionan servicios novedosos tales como gestión centralizada de contenidos multimedia [1]. Podemos acceder a más servicios pero esto lleva consigo el manejo de nuevos dispositivos, obteniendo disponibilidad a cambio de complejidad de uso.

En este contexto aparecen sistemas que persiguen la personalización del entorno. Un ejemplo de este tipo puede verse en dispositivos móviles sensibles a la localización, capaces de adaptarse al lugar al que accede el usuario, o en cualquier sistema que ajuste su comportamiento a las preferencias de distintos usuarios cercanos. Una de las principales líneas de investigación de nuestro grupo es la personalización de dispositivos móviles y de electrodomésticos tradicionales, aportando inteligencia al hogar [2]. Para alcanzar nuestro objetivo proponemos el uso de sistemas multiagente, ya que son idóneos para el desarrollo de sistemas distribuidos, inteligentes y autónomos. El trabajo que aquí proponemos continúa con lo expuesto en [3].

Con nuestro sistema hemos convertido un dispositivo móvil en un mando a distancia universal capaz de adaptarse a la localización de su usuario, proporcionando los interfaces correspondientes a los servicios disponibles en su localización. Además hemos implementado un servicio multimedia que permite que los contenidos solicitados por los usuarios sigan los movimientos de sus solicitantes dentro de la vivienda en tiempo real, de forma que no exista pérdida de información. Los agentes inteligentes proporcionan la tecnología necesaria para alcanzar el grado de distribución, autonomía e inteligencia requerido.

El documento sigue la siguiente estructura: se realiza un resumen del estado del arte en la sección 2, y se detalla nuestro hogar inteligente en las secciones 3, 4 y 5. En ellas se especifica la arquitectura de agentes, la funcionalidad de los agentes y las interfaces entre ellos, y se introduce la ontología utilizada. Finalmente, se extraen conclusiones y se trata el trabajo futuro en la sección 6.

2 Hogares Inteligentes

Podemos definir un entorno inteligente como *aquel capaz de adquirir y aplicar conocimiento sobre sus habitantes y sus alrededores para adaptarse a ellos y conocer los objetivos de confort y eficiencia* [4]. Dichos objetivos suelen centrarse en la adaptación del entorno a las preferencias de los usuarios, para incrementar su rendimiento en las tareas diarias, y para op-

timizar el consumo de energía de los sistemas implicados.

2.1 Personalización de servicios en hogares inteligentes

La personalización de sistemas y servicios no es un tema novedoso. De hecho, la mayoría de las líneas de investigación acerca de agentes software de usuario se centran en la posibilidad de configurar un sistema software de acuerdo a las preferencias del usuario. Puede existir un agente asociado a cada usuario, como se describe en [5], o un único agente que sirve a todos los usuarios que acceden al sistema, como en [6]. Aunque la personalización de servicios es factible para cualquier entorno de usuario, no hay duda de que los hogares inteligentes son especialmente adecuados para ello, ya que en el hogar los usuarios son más proclives al uso de sistemas diseñados para ofrecer confort y facilidad de uso.

El Libro Blanco del Hogar Digital, editado por Telefónica [7] proporciona una taxonomía para servicios disponibles en el hogar digital, que pueden personalizarse en gran medida para ajustarse a las preferencias de los usuarios. Por ejemplo, podemos pensar en un hogar en el que las luces se apagan y se encienden siguiendo al usuario por la vivienda, en el que una llamada telefónica se convierte automáticamente en una videoconferencia si el usuario tiene un televisor o una pantalla de ordenador cerca, o en el que el televisor emite la totalidad o parte de los contenidos, en función de la edad de los televidentes.

2.2 Hogares inteligentes y sistemas multiagente

Para alcanzar los objetivos que estamos planteando, un sistema domótico se compone de dos conjuntos de dispositivos: unos recogen información del entorno en el que están situados -sensores- y los restantes son capaces de actuar sobre las condiciones del entorno -actuadores-. El sistema procesará los datos recogidos por los sensores y, de acuerdo a los objetivos establecidos, empleará a los actuadores para alterar el entorno del usuario. La complejidad de la domótica estriba en las decisiones que toma el sistema acerca de las acciones necesarias. Estas acciones se basan en la información recogida por los sensores. La obtención de información deberá realizarse de forma distribuida, autónoma e inteligente, lo que sugiere el uso de agentes software para el desarrollo de este tipo de sistemas.

Existen diferentes definiciones para el concepto de agente software. Desde el punto de vista del diseño y la tecnología, podemos decir que un agente software es un programa auto-contenido capaz de controlar su propia toma de decisiones, y actuar basándose en su percepción del entorno que le rodea, y persiguiendo uno o más objetivos [8]. Desde un punto de vista funcional o desde la perspectiva de un usuario, un agente software puede verse como una entidad software en la que pueden delegarse tareas [9]. La última definición, aunque más sencilla, sugiere más claramente que esta tecnología puede solucionar el problema de la automatización inteligente del entorno.

3 Una plataforma de agentes de hogar inteligente

Una primera contribución de este trabajo es una arquitectura para la construcción de un hogar digital basado en agentes, que hemos llamado iHAP (*intelligent Home Agent Platform*). Esta arquitectura cuenta con un conjunto de dispositivos distribuidos en el entorno. De acuerdo al grado de autonomía e inteligencia de los dispositivos, principalmente determinado por su capacidad computacional para albergar agentes, podemos dividirlos en cuatro grupos [10]:

- **iHAP Central System (iHAP o CS)**. Está ubicado en la pasarela residencial [11] y contiene la plataforma de agentes que soporta la existencia de todos los agentes de la vivienda. Contiene aquellos agentes que actúan a nivel global en la vivienda, y no están asociados a un sensor o actuador específicos o a una localización o habitación determinadas. En el iHAP Central System se encuentran también los agentes utilizados para el control de dispositivos no inteligentes, es decir, dispositivos domóticos sin suficiente capacidad de proceso para alojar sus propios agentes. Aunque la fiabilidad del iHAP Central System es esencial para el funcionamiento adecuado del sistema, en caso de fallo en el CS, los agentes que están distribuidos en el hogar entrarían en un modo seguro, en el que se les permite ofrecer una funcionalidad básica.

- **Dispositivos personales**. Cada usuario porta un dispositivo móvil -teléfono móvil, PDA- que contiene los agentes necesarios para identificar al usuario en el sistema, determinar la ubicación del usuario en el hogar, y mostrar los interfaces apropiados a los servicios disponibles cuando sea necesario.

- **Dispositivos con agentes**. Son dispositivos sensores y actuadores con cierto grado de autonomía, normalmente procedente de los agentes activos sobre una máquina virtual Java empotrada.

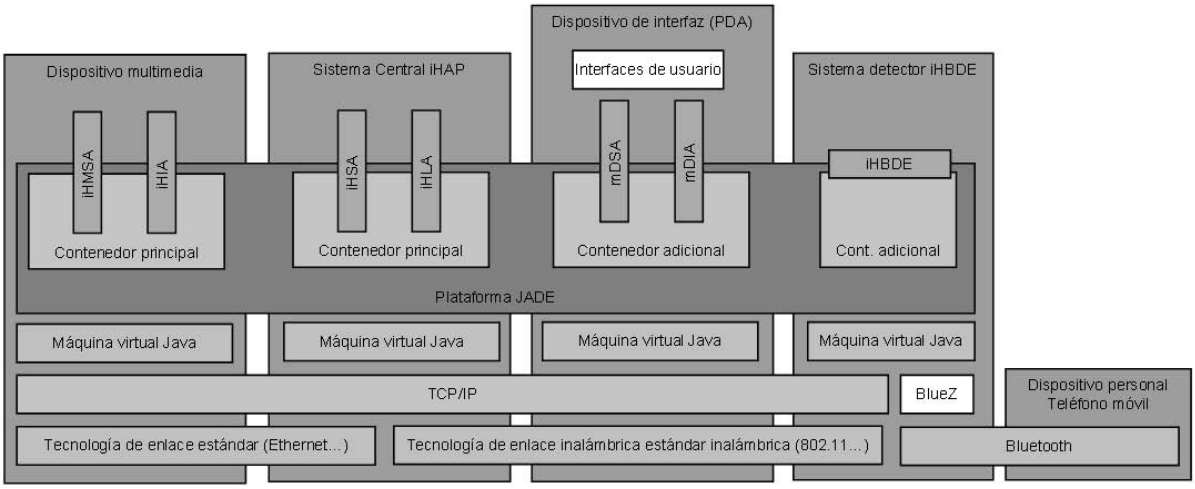


Figura 1: Arquitectura del sistema

- **Dispositivos sin agentes**. Están controlados por el CS, ya que no tienen la capacidad necesaria para funcionar de forma autónoma. Están conectados a él mediante tecnologías de bus estándar. El resto de dispositivos, con agentes y personales, se comunican entre sí y con el iHAP CS a través de TCP/IP. Aunque sería posible utilizar cualquier tecnología de enlace, hemos optado por el uso de protocolos de comunicación inalámbrica -Bluetooth, WLAN-.

En la Fig. 1 se puede ver la arquitectura del sistema en la que se muestran los distintos niveles de comunicación. El entorno de desarrollo de libre distribución utilizado ha sido la plataforma de agentes JADE (*Java Agent DEvelopment framework*) [12], que nos libera de la realización de tareas de bajo nivel relacionadas con el ciclo de vida de los agentes y el intercambio de mensajes. La plataforma JADE se extiende a todos los elementos del sistema que contienen agentes, éstos viven dentro de un contenedor asociado a cada elemento, y el contenedor principal de la plataforma reside en el iHAP CS. Por otra parte, el uso de las especificaciones FIPA (*Foundation for Intelligent Physical Agents*) [13], garantizan cierto grado de interoperabilidad con otros sistemas basados en agentes. El uso de Java asegura la portabilidad del código a otras máquinas distintas.

Las interfaces tanto entre un dispositivo con agentes y su actuador o sensor asociado, como con el resto de elementos del sistema se proporciona mediante un sistema empotrado a través de una máquina virtual Java. En estos momentos estamos trabajando con diferentes tarjetas, tratando de encontrar la opción más conveniente para cada dispositivo. Una de las opciones es utilizar tarjetas TINI (*Tiny InterNet Interface*) [14].

En esta primera implementación, hemos contemplado la existencia de dos servicios en nuestro hogar inteligente: distribución de audio/video multiroom e interfaces de usuario sensibles a la localización. Con el primero de ellos los usuarios tienen acceso a contenidos multimedia desde cualquier localización en el hogar digital, a través de un repositorio de contenido centralizado, normalmente ubicado en la pasarela residencial, y se permite que los contenidos multimedia sigan al usuario si decide cambiar la habitación en la que esta disfrutando de los mismos. Con el segundo, los usuarios interactúan con un interfaz genérico que se adapta a los servicios multimedia disponibles en una localización concreta, lo que evita que el usuario tenga que aprender distintos interfaces.

4 Arquitectura de agentes

La Fig. 1 se muestra la arquitectura del iHAP particularizada para los servicios descritos. Podemos ver los distintos elementos del sistema y los agentes software que residen en ellos, que confieren la funcionalidad requerida.

4.1 Dispositivo móvil

La funcionalidad asociada a este dispositivo se ha dividido en dos bloques separados. Por una parte, un dispositivo móvil de interfaz, en este caso una PDA con una tarjeta 802.11 que puede usarse a modo de mando a distancia para interactuar con el sistema. El interfaz mostrado se adapta a los servicios disponibles, dependiendo de la localización del usuario. Por otra, un dispositivo móvil personal, que se refiere a un teléfono móvil con interfaz Bluetooth que permite al sistema identificar al usuario y determinar su localización en el hogar.

Aunque podríamos haber utilizado un único dispositivo, hemos considerado más interesante el uso de dos, ya que la separación de funcionalidad permite utilizar un único mando a distancia 802.11 para el hogar, mientras identificamos a los usuarios a través de sus teléfonos móviles Bluetooth. El dispositivo móvil personal no contiene agentes, mientras que el dispositivo

móvil de interfaz contiene dos agentes:

- **mDSA** (*mobile Device Service Agent*). Este agente se encarga de mostrar en la pantalla de la PDA interfaces a partir de los que el usuario podrá manejar los servicios. Del agente iHSA recibe mensajes cuyo contenido hace referencia al interfaz correspondiente.

- **mDIA** (*mobile Device Interface Agent*). La función principal de este agente es la interacción con el usuario, y recibe sus peticiones a través de los interfaces visibles en la PDA. Una vez recibida la orden, envía una solicitud de servicio al agente iHSA, que se encarga de llevar a cabo la petición.

4.2 Equipo de Detección Bluetooth (BDE)

Se refiere a un dispositivo empotrado provisto de un interfaz Bluetooth que se conecta con el dispositivo personal del usuario, y realiza medidas de potencia para estimar su localización. Esta funcionalidad está soportada mediante el siguiente agente:

- **iHBDE** (*intelligent Home Bluetooth Detection Equipement*). Todos los agentes de los BDEs de cada localización del hogar perciben al usuario mediante la potencia y la intensidad de la señal recibida. Si esos parámetros superan cierto umbral, se envía un mensaje al agente iHLA, informando del grado de cercanía del usuario.

4.3 Dispositivo multimedia

Se trata de un dispositivo reproductor estándar de audio/video con cierto grado de autonomía proporcionada por los siguientes agentes:

- **iHMSA** (*intelligent Home Multimedia Service Agent*). Su misión es controlar el servicio multimedia que se ofrece en el dispositivo en el que reside. Recibe órdenes del agente iHSA para que actúe sobre el estado del servicio que controla. En caso de que el servicio finalice su ejecución informa al iHSA.

- **iHMIA** (*intelligent Home Multimedia Interface Agent*). El dispositivo multimedia debe ser capaz de mostrar un interfaz que permite la interacción con el usuario. Este agente recibe las peticiones de actuación de los usuarios y las reenvía al agente iHSA.

4.4 iHAP Central System

En esta implementación contiene dos agentes:

- **iHLA** (*intelligent Home Location Agent*). El objetivo de este agente es determinar la localización de los usuarios, e informar de ello al agente iHSA. La decisión se tomará de forma centralizada a partir de la información proporcionada por los agentes iHBDE de cada ubicación.

- **iHSA** (*intelligent Home Service Agent*). Almacena toda la información referente a los usuarios, como es su localización en el hogar inteligente, los servicios a los que tiene acceso cada usuario en cada habitación, referencias a los distintos interfaces, el estado de todos los servicios de cada localización y el estado de los servicios que solicita cada usuario. Recibe información acerca de la localización de los usuarios, y emite dos tipos de órdenes. Si un usuario sale de una localización y tiene servicios activos en ella, ordena parar los servicios, y almacena el estado de los mismos. En una nueva localización en la que entra un usuario, este agente solicita la ejecución de aquellos servicios activos en la localización anterior, en caso de que el usuario viniera de otra localización. Además, recibe las peticiones de actuación sobre los servicios, emitidas por los agentes de interfaces, y ordena la actuación sobre el servicio que corresponda. Por último, es informado cuando un servicio finaliza su ejecución, y actualiza la información que mantiene sobre ese servicio.

5 Comunicación entre agentes

Con el fin de garantizar la interoperabilidad, la comunicación entre los agentes se realiza utilizando mensajes y actos comunicactivos. Se utilizan mensajes ACL (*Agent Communication Language*) y actos comunicativos definidos por la organización FIPA en las especificaciones FIPA ACL [15] y [16], respectivamente.

Describiremos el comportamiento del sistema iHAP de acuerdo a los servicios de esta primera implementación, que considera un sencillo hogar con un salón y un dormitorio conectados por un pasillo. En la Fig. 2 podemos ver que en cada habitación existe un Equipo de Detección Bluetooth y un Dispositivo Multimedia. El iHAP Central System puede situarse en cualquier ubicación dentro del hogar.

5.1 Proceso de detección de usuarios y seguimiento de servicios

Inicialmente, una vez arrancada la plataforma, el sistema queda a la espera de la detección de usuarios. Los agentes iHBDE perciben a los usuarios mediante Bluetooth, y cuando la potencia e intensidad recibidas rebasan cierto umbral, envían un mensaje (**M1**) al agente iHLA informándole de la presencia del usuario.

Cuando el agente iHLA recibe el mensaje, busca la información almacenada del usuario, y compara todos los informes recibidos asociados a las distintas localizaciones, decidiendo en qué localización se encuentra el usuario, y enviando un mensaje al agente iHSA (**M2**).

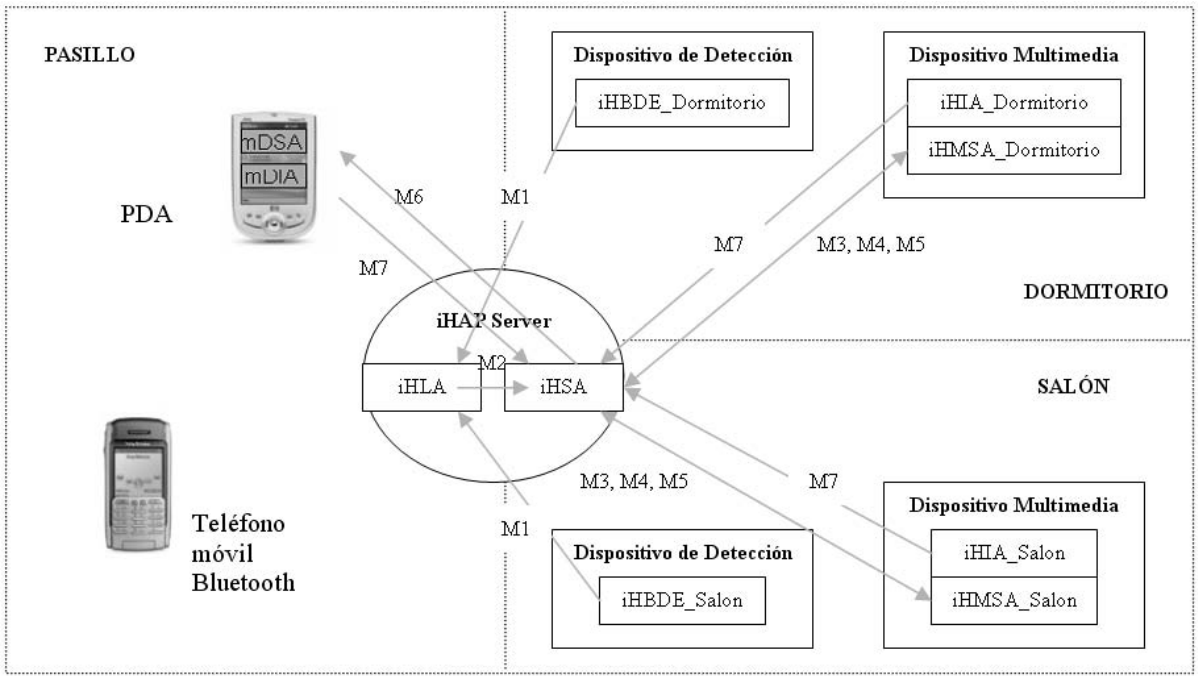


Figura 2: Escenario de aplicación

El agente iHSA marca como presente al usuario en la nueva localización, realiza una búsqueda de servicios activos asociados al usuario en la antigua localización, si esta existe, y ordena al agente iHMSA de la antigua localización que los pare (**M3**). A continuación, se devuelve el estado en el que se ha detenido cada servicio (**M4**) al agente iHSA, que lo almacena para su uso futuro.

El agente iHSA comprueba la existencia de los servicios parados en la nueva localización, y ordena al agente iHMSA de la nueva localización, que los arranque según el estado en el que fueron parados en la anterior localización (**M5**).

En caso de que los servicios no estén disponibles, se almacena su estado para que puedan ser ejecutados en posteriores localizaciones accedidas por el usuario.

Por último, el agente iHSA envía al agente mDSA una referencia al interfaz que debe mostrar en la pantalla de la PDA del usuario (**M6**). Este interfaz indica al usuario dónde se encuentra, si tiene servicios activos, y proporciona acceso a los servicios disponibles en la localización. De esta forma el estado de los servicios se conserva de una a otra habitación, siguiendo los movimientos de los usuarios que los han solicitado.

5.2 Proceso de petición de servicios

Cuando un usuario es detectado en una localización, aparece un interfaz en su PDA que le permite controlar los servicios disponibles en esa localización. Además existen interfaces estáticos en los dispositivos multimedia para controlar el servicio que ofrecen, aunque ningún usuario haya sido detectado en la localización. Esto nos permite manejar los servicios en ausencia de dispositivos móviles. En ambos casos la comunicación entre los agentes implicados es la misma; la única diferencia son los agentes que interactúan con los usuarios.

Si un usuario solicita un servicio desde la PDA el agente mDIA asociado al usuario envía una petición de servicio al agente iHSA (**M7**), que tras actualizar el estado del servicio en los datos que mantiene, encamina la petición hacia el agente iHMSA de la localización en la que se encuentre el usuario (**M3** ó **M5**).

En caso de que el usuario pause el servicio, el agente iHSA recibe el estado del servicio (**M4**), para poder volver a ejecutarlo posteriormente.

Si el servicio se solicita desde alguno de los dispositivos multimedia, el agente iHIA presente en el dispositivo envía una petición de servicio al agente iHSA (**M7**) que, tras actualizar el estado del servicio en los datos que mantiene, encamina la petición hacia el agente iHMSA de la localización de la que proceda la petición (**M3** ó **M5**).

En caso de que el usuario pause el servicio, el agente iHSA recibe el estado del servicio (**M4**), para poder volver a ejecutarlo posteriormente.

El sistema permite que usuarios sin dispositivos móviles registrados puedan acceder a los servicios que se ofrecen, pero sin valor añadido, es decir, sin seguimiento en el hogar.

Por último, cuando alguno de los servicios que controlan los agentes iHMSA finaliza su ejecución, el agente iHMSA correspondiente informa al agente iHSA para que actualice la información que mantiene (**M4**).

Cuadro 1: Ontolología soportada por el iHAP

Estructura	Parámetro	Descripción	Tipo	Rango
`User`	`:identity`	Este tipo de objeto representa la identidad de un usuario.	`String`	
`Location`	`:name`	Este tipo de objeto representa una localización del hogar inteligente.	`String`	`salon, dormitorio, pasillo`
`Power`	`:value`	Este tipo de objeto representa la potencia con la que se percibe a un usuario en alguna localización del hogar inteligente.	`Integer`	
`Service`	`:name`	Este tipo de objeto representa un servicio multimedia que se ofrece en el sistema.	`String`	`audio, video`
`State`	`:value`	Este tipo de objeto representa el estado de un servicio	`String`	
`Action`	`:name`	Este tipo de objeto representa el tipo de acción a realizar sobre un servicio.	`String`	`play, pause, stop`
`Interface`	`:name`	Este tipo de objeto representa una referencia al interfaz que debe mostrarse al usuario.	`String`	
`Notice`: Este objeto representa un predicado que expresa que un usuario ha sido percibido desde una localización, con un determinado valor de potencia.	`:user`	Usuario detectado	`User`	
	`:location`	Localización desde la que ha sido detectado	`Location`	
	`:power`	Potencia con que ha sido detectado	`Power`	
`Detect`: Este objeto representa un predicado que expresa que un usuario ha sido detectado en una localización.	`:user`	Usuario detectado	`User`	
	`:location`	Localización en la que está presente	`Location`	
`Command`: Este objeto representa la acción de solicitar un servicio.	`:service`	Servicio sobre el que se va a actuar	`Service`	
	`:action`	Acción a realizar	`Action`	
	`:state`	Estado que debe tener el servicio	`State`	
`Report`: Este objeto representa un predicado que informa del estado de un servicio.	`:service`	Servicio del que se informa	`Service`	
	`:demandedState`	Estado del servicio	`State`	
`Show`: Este objeto representa una acción en la que se solicita que se muestre un interfaz.	`:interface`	Referencia al interfaz que debe mostrarse	`Interface`	

Cuadro 2: Mensajes utilizados en la comunicación

Mensaje	Performative	Sender	Receiver	Content	Language	Ontology
M1: **detectionReport**	inform	El agent-identifier de alguno de los agentes iHBDE	El agent-identifier del agente iHLA	Una estructura Notice	FIPA-SL	iHAP-Ontology
M2: **userLocation**	inform	El agent-identifier del agente iHLA	El agent-identifier del agente iHSA	Una estructura Detect	FIPA-SL	iHAP-Ontology
M3: **serviceAction**	request	El agent-identifier del agente iHSA	El agent-identifier de algún agente iHMSA	Una estructura Command	FIPA-SL	iHAP-Ontology
M4: **stateService**	inform	El agent-identifier de algún agente iHMSA de alguna localización	El agent-identifier del agente iHSA.	Una estructura Report	FIPA-SL	iHAP-Ontology
M5: **startService**	request	El agent-identifier del agente iHSA	El agent-identifier de algún agente iHMSA	Una estructura Command	FIPA-SL	iHAP-Ontology
M6: **showInterface**	request	El agent-identifier del agente iHSA	El agent-identifier de algún agente mDSA	Una estructura Show	FIPA-SL	iHAP-Ontology
M7: **requestService**	request	El agent-identifier de algún agente de interfaz (mDIA o iHIA)	El agent-identifier del agente iHSA	Una estructura Command	FIPA-SL	iHAP-Ontology

5.3 Introducción a la ontología soportada por la plataforma iHAP

Para garantizar el correcto entendimiento entre diferentes agentes, FIPA se apoya en los conceptos de lenguaje y ontología. El lenguaje define las reglas con las que los diferentes elementos del contenido de un mensaje se pueden combinar. La ontología describe los elementos que pueden utilizarse como contenido de un mensaje [17].

Una ontología define un vocabulario de términos y unas relaciones entre los términos de ese vocabulario. Dichas relaciones pueden ser estructurales o semánticas. Las relaciones estructurales describen cómo algunos elementos están definidos mediante otros elementos, de un modo similar a los datos miembro de un objeto software. Las relaciones semánticas se refieren a que algunos elementos extienden o concretan otros elementos, de un modo similar a lo que ocurre en la herencia en programación orientada a objetos.

Los diferentes agentes que componen nuestro sistema distribuido van a poder compartir información mediante la utilización de ontologías, ya que su uso permite una representación uniforme de la información. Por tanto, los agentes van a entenderse y la adición de nuevos agentes es más sencilla, de forma que se facilita la escalabilidad del sistema.

La comunicación entre los agentes de la plataforma iHAP se apoya en la utilización de la ontología *FIPA-Agent-Management* para los conceptos básicos de la plataforma de agentes [18], y en la ontología *iHAP-Ontology* para el diálogo entre agentes.

La ontología *FIPA-Agent-Management* permite identificar la plataforma y todos los agentes presentes en ella, incluyendo aquellos que pertenecen a la plataforma JADE, como es el agente DF (*Directory Facilitator*). También es posible describir los servicios que los agentes registran en el DF, y se define la interacción entre el DF y el resto de agentes.

En el Cuadro 1 se describen algunos de los conceptos, predicados y acciones de la ontología *iHAP-Ontology* utilizados en las interacciones entre los agentes de la plataforma iHAP. Los objetos de la ontología *iHAP-Ontology* que se han mostrado complementan a la ontología *FIPA-Agent-Management*, consiguiendo que el diálogo entre agentes sea posible, y que la representación de la información que se utiliza quede definida de forma homogénea. Mediante esta ontología se definen los objetos necesarios en la prestación de servicios audio y video multiroom, e interfaces de acceso sensibles a la localización.

5.4 Definición de los mensajes intercambiados

Los mensajes utilizados en la interacción entre los agentes, detallados en el Cuadro 2, se implementan con mensajes FIPA-ACL, cuya sintaxis completa hemos omitido por razones de espacio.

6 Conclusiones y trabajo futuro

Este documento proporciona una infraestructura para desarrollar entornos inteligentes mediante agentes software. Define claramente la funcionalidad de cada uno de los agentes que componen el sistema, las interfaces entre ellos, y aquellos aspectos de la ontología *iHAP-Ontology* necesarios para la comunicación entre los agentes. La principal ventaja de la utilización de agentes es la autonomía que pueden ofrecer. Un agente inteligente basa su comportamiento en un conjunto de objetivos de alto nivel, y determina de forma autónoma las acciones necesarias para cumplir dichos objetivos. Estas acciones pueden incluir la interacción y cooperación con otros usuarios. De hecho, los sistemas multiagente son sistemas distribuidos muy apropiados para su aplicación en entornos de hogar inteligente, donde es necesaria la coordinación de sensores y actuadores distribuidos. Los dos servicios que hemos implementado sobre la arquitectura propuesta, muestran como los agentes software pueden ayudar a adaptar el entorno a las preferencias y deseos del usuario.

Actualmente estamos trabajando en la mejora de los Equipos de Detección Bluetooth, haciendo el proceso de localización más rápido, más fiable y más eficiente en términos de consumo de energía. Se está desarrollando una arquitectura jerárquica distribuida para un Sistema de Detección de Presencia (iHPDS) que extiende la funcionalidad de los BDEs y del agente iHLA. Además, estudiamos la posibilidad de interacción entre usuarios y el sistema mediante la integración del dispositivo personal y el de interfaz en un único teléfono móvil. Por otra parte, se están desarrollando e implementando nuevos servicios, e interfaces para gestionar remotamente el hogar digital diseñado.

Finalmente, estamos trabajando en la ontología *iHAP-Ontology* para ampliar su uso en el sistema completo, no únicamente en el apoyo a la comunicación entre agentes. Esta extensión abarca el modelado de conceptos relacionados con los entornos inteligentes en general, y se prevé su utilización en la toma de decisiones de los agentes.

Agradecimientos

Este trabajo se ha realizado gracias al proyecto MCYT-TIC2003-09192-C11-05 del Ministerio de Ciencia y Tecnología, y al proyecto UAH-PI-2003/001 de la Universidad de Alcalá.

Referencias

[1] PROHOME. “Proyecto mercahome, informe b1. Análisis de la oferta actual” Tech. Rep., Diciembre 2004. [Online]. Disponible: www.casadomo.com/prohome/

[2] Miguel A. Lopez, Alvaro Paricio, Juan R. Velasco e Iván Marsá. “Arquitectura de agentes para entornos domóticos” en XIV Jornadas Telecom I+D. 2004.

[3] Juan R. Velasco, Ivan Marsá, Andrés Navarro, Miguel A. López, Antonio J. Vicente, Enrique de la Hoz, Alvaro Paricio y Miriam Machuca. “Personalización de servicios multimedia en el hogar digital inteligente”. Actas del X Congreso Internet, Telecomunicaciones y Sociedad de la Información (Mundo Internet 2005), pp. 467-474.

[4] D. J. Cook y M. Youngblood. “Smart Homes”. Beckshire Encyclopedia of Human-Computer Interaction. Berkshire Publishing Group, 2004.

[5] H. Müller, T. Hilbrich, y R. Kühnel. “An assistant agent”. Fundamenta Informaticae, no. 34, pp. 1-10. 1999.

[6] M. Pazzani and D. Billsus. “Adaptive web site agents”. Autonomous Agents and Multiagent Systems, no. 5, pp. 205218, 2002.

[7] “Libro blanco del hogar digital y las infraestructuras comunes de telecomunicaciones”. Telefónica. Tech. Rep. 2003. [Online] Disponible: www.telefonica.es/index/libroblancohogardigital.html

[8] N. Jennings and M. Wooldridge. “Software agents”. IEE Review, pp. 1720, January 1996.

[9] P. Janca, “Pragmatic application of information agents”. BIS Strategic Decisions. Tech. Rep., 1995.

[10] Miguel A. Lopez, Iván Marsá Maestre, Andrés Navarro y Juan R. Velasco. “Arquitectura para un sistema domótico basado en agentes”. Conferência Ibero-Americana IADIS WWW/Internet, pp. 469-472, 2004.

[11] D. Valtchev and I. Frankov. “Service gateway architecture for a smart home”. IEEE Communications Magazine, pp. 126-132, April 2002.

[12] CSELT, JADE homepage: http://jade.cselt.it.

[13] FIPA, FIPA homepage, http://www.fipa.org/.

[14] “Tiny internet interface”. [Online]. Disponible: http://www.ibutton.com/TINI/index.html

[15] FIPA, “FIPA Acl Message Structure Specification”. Document SC00061GX, 2002. [Online]. Disponible : http://www.fipa.org

[16] FIPA, “FIPA Communicative Act Library Specification”. Document SC00037J, 2002. [Online]. Disponible: http://www.fipa.org

[17] G. Caire, David Cabanillas. “JADE tutorial application-defined content languages and ontologies”. TILAB S.p.A. 2002. [Online]. Disponible: http://jade.cselt.it

[18] FIPA, “FIPA Agent Management Specification”. Document SC00023J, 2002. [Online]. Disponible: http://www.fipa.org

Localización de Usuarios en Interiores en Redes Móviles de Tercera Generación

F. Gil Castiñeira, F. J. González Castaño, J. M. Pousada Carballo, P.S. Rodríguez Hernández
Departamento de Ingeniería Telemática. Universidad de Vigo
ETSE de Telecomunicación. C/ Maxwell S/N. Campus Universitario.
36310 - Vigo (Pontevedra)
E-mail: {xil,javier,chema,pedro}@det.uvigo.es

Abstract

In this paper we propose the integration of high-precision indoor location networks –such as RADAR or Bluetooth Location Networks (BLN)– into the 3rd Generation Mobile Networks architecture. Thus, in our approach, such networks become Assistant Location Networks*, to improve Cell-ID location precision indoors. We prove that Assistant Location Networks are scalable. We propose their integration into the core network via a NBAP extension by simulating the overload generated in a Manhattan Grid Model. We prove that the impact on the traffic of such integration is admissible in practice even in densely populated scenarios.*

1. Introducción

Telefónica Móviles lanzó en 2002 el *Programa de Promoción Tecnológica del UMTS* (PPT-UMTS), mediante el que se convocaba a la Universidad a realizar propuestas de I+D+I, con el objeto de identificar nichos dentro de la tecnología UMTS (Universal Mobile Telecommunications System) que pudieran llevar a investigaciones orientadas al mercado. A nuestro equipo le fue concedido uno de los cinco proyectos PPT-UMTS de 2002, por las ideas y los resultados presentados en este escrito. En él, presentamos una *contribución conceptual* –la integración de UMTS y redes auxiliares de localización no-3G "externas"– y la correspondiente solución tecnológica.

En las redes móviles 3G, las Aplicaciones Basadas en Localización (LBA, Location Based Application) están soportadas por *servicios de localización* estándar [1]. Sin embargo, dichos servicios presentan varios inconvenientes, ya que la precisión de la localización mediante identificadores de celda de UMTS (Cell-ID) está limitada por el tamaño de celda. OTDOA (Observed Time Difference of Arrival) y AGPS (Assisted GPS) son más precisos, pero sólo son fiables en exteriores.

En este trabajo proponemos la integración en la arquitectura de UMTS de redes de localización en interiores de alta precisión, tales como RADAR [2], de Microsoft, o BLN (Bluetooth Location Network) [3]. Así, en nuestro planteamiento, una red dada de localización en interiores se convierte en una Red Auxiliar de Localización (RAL), con mayor precisión que mediante Cell-ID. Demostramos que las RALs son escalables y proponemos su integración en el núcleo de red UMTS.

Trabajos previos han propuesto muchos sistemas de posicionamiento de usuarios en interiores. Algunos de ellos se basan en dispositivos especializados, no comerciales [4, 5, 6, 7]. Otras soluciones se basan en equipos de usuario disponibles comercialmente. En los sistemas de posicionamiento de usuarios en interiores, un *servidor de localización* sigue la pista a los *badges* de los usuarios. Los *badges* pueden ser dispositivos independientes (e.g. tags RFID) o empotrados (e.g. modems Bluetooth en equipos de usuario comerciales). Una infraestructura de *nodos de detección* cubre el escenario de aplicación, interactuando con los badges de los usuarios. Los posibles nodos de detección pueden ser detectores RFID, nodos estáticos Bluetooth, receptores de ultrasonidos, antenas de RF direccionales... El rango de cada nodo de detección define una *picocelda de localización*. Éstas envían la información que recopilan hacia el servidor de localización a través de una LAN cableada o una red ad-hoc. A partir de esa información, dicho servidor de localización puede calcular la posición de los usuarios.

Nuestra propuesta de integración se basa en la observación de que el servidor de localización es típicamente un nodo de Internet. Por lo tanto, es fácil integrarlo en la arquitectura UMTS.

Nuestro análisis de escalabilidad se basa en la evaluación de la tasa de tráfico de las LBA en el núcleo de red UMTS, para estimar su impacto, caso de que lo haya, sobre el tráfico de voz+datos. En un escenario interior cubierto por *microceldas* UMTS tienen lugar frecuentes cambios de celda, que causan señalización de localización. El despliegue de una RAL puede verse como la definición de picoceldas de localización (*picoceldas RAL*) dentro de las microceldas UMTS. Las picoceldas RAL son transparentes a las llamadas de UMTS (manejadas por las microceldas UMTS). Sin embargo, cada vez

que un equipo de usuario activo cruza la frontera de una picocelda RAL, se transmite un mensaje de localización al núcleo de red. Como resultado, el tráfico de localización crece con la precisión de la localización. No obstante, como veremos en la sección 3, el tráfico de localización es relativamente bajo, incluso en escenarios densamente poblados con picoceldas RAL de 10 m de radio.

El resto de este trabajo se organiza como sigue: en la sección 2 describimos sucintamente la arquitectura de localización de UMTS. En la sección 3 presentamos nuestro análisis de la escalabilidad de RAL. La sección 4 se ocupa de la integración RAL-UMTS. Finalmente, la sección 5 muestra las conclusiones.

2. La arquitectura de localización de UMTS

La Fig. 1 muestra los elementos de red cuyas entidades funcionales dan soporte a los servicios de localización.

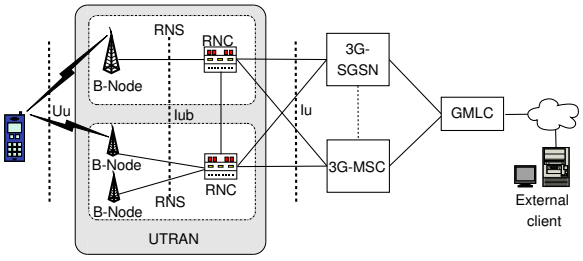


Figura 1: Elementos de red para los servicios de localización

El Centro de Localización de Móviles con funciones de Pasarela (GMLC, Gateway Mobile Location Center) es un componente fundamental en el sistema de localización. Es el primer nodo al que las aplicaciones externas acceden cuando solicitan posicionamiento de equipos de usuario. El GMLC solicita información de localización a los otros elementos de red –Nodo Servidor de Soporte GPRS (SGSN, Serving GPRS Support Node), Central de Conmutación Móvil (MSC, Mobile Services Switching Center)– y retorna las posiciones de los equipos de usuario a las aplicaciones externas. El MSC y el SGSN contienen información relativa a las autorizaciones/suscripciones del sistema de localización.

2.1. Localización UMTS en entornos microcelulares

Obviamente, la precisión mediante Cell-ID es relativamente alta en entornos microcelulares debido al pequeño radio de las celdas. La mayores tasas de tráfico de localización ocurren cuando las LBAs siguen las posiciones de los equipos de usuario en tiempo real. Mediante Cell-ID, esto es equivalente a seguir la pista de los cambios de celda.

El estándar UMTS define un método para enviar informes de localización siempre que el equipo de usuario se incorpore a una celda. Hay dos alternativas:

1. El equipo de usuario puede enviar informes de localización, como se muestra en la Fig. 2.
2. La Red de Acceso Radio Terrestre UMTS (UTRAN, Universal Terrestrial Radio Access Network) puede enviar los informes. Con este objeto, el protocolo de la parte de aplicación RAN (RANAP, Radio Access Network Application Part) [9] tiene un mensaje `LOCATION_REPORTING_CONTROL`, con los siguientes campos:
   - *Request type:* determina si el proceso de localización debe empezar o parar y si la respuesta debe ser inmediata o seguir a un evento (cambio de área de localización).
   - *Response information:* identificador y coordenadas del área de localización.

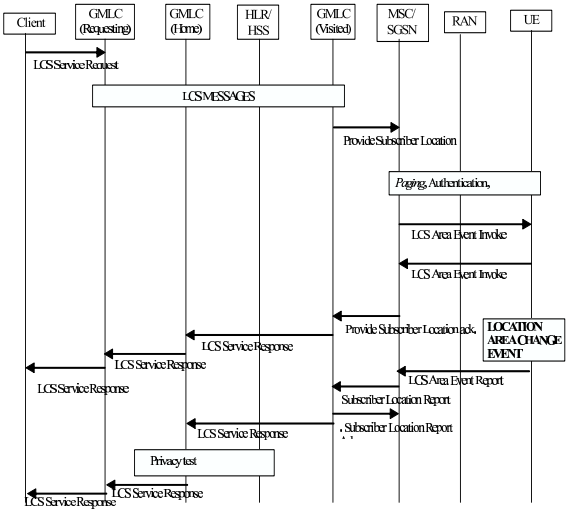


Figura 2: Notificación del cambio de área de localización

Es posible lanzar un evento de localización siempre que el equipo de usuario entre en una nueva área de localización. En nuestro caso, las áreas de localización son celdas, y el tamaño de celda determina la precisión de la localización mediante Cell-ID.

2.2. Uso del ancho de banda en los mensajes de localización

Trabajos previos han caracterizado el tráfico UMTS teniendo en cuenta en diferentes objetivos de investigación [10, 11]. En [10], García et al evaluaron el ancho de banda del tráfico de llamadas a través de la interfaz *Iu*, entre un Nodo B y su Controlador de Red Radio (RNC, Radio Network Controller). En nuestro caso, queremos evaluar el tráfico *Iu* generado por las operaciones de localización. El interfaz *Iu* agrega los tráficos existentes de la interfaz *Uu* (equipo de usuario-UTRAN).

Las solicitudes de localización de equipos de usuario pueden remitirse directamente al GMLC por medio del Protocolo de Localización en Redes Móviles (MLP, Mobile Location Protocol) [8] definido en XML (eXtensible Markup Language), a través de HTTP, WSP, SOAP, etc. En el caso de HTTP, la respuesta a una solicitud periódica de localización (caja inferior) tiene una longitud de 685 bytes.

```
POST /newResult HTTP/1.1
Host: terminal.movil.com
Content-Type: text/xml
Content-Length: 588
<?xml version="1.0"?>
<!DOCTYPE svc_init SYSTEM "
    MLP_SVC_RESULT_300.DTD">
<svc_result ver="3.0.0">
<hdr ver="3.0.0">
<client>
<id>aUser</id>
<pwd>aPwd</pwd>
</client>
</hdr>
<tlrep ver="3.0.0">
  <req_id>25267</req_id>
  <trl_pos trl_trigger="PERIODIC">
    <msid>461011678298</msid>
    <pd>
      <time utc_off="+0300"
          >20020813010423</time>
      <shape>
        <CircularArea srsName="www.epsg
            .org#4326">
          <coord>
            <X>35 35 24.139N</X>
            <Y>139 35 24.754E</Y>
          </coord>
          <radius>15</radius>
        </CircularArea>
      </shape>
    </pd>
  </trl_pos>
  <time_remaining>00010000</
      time_remaining>
</tlrep>
</svc_init>
```

Alternativamente, la solicitud se podría transmitir al MSC, de donde posteriormente el GMLC tomaría los datos de localización. En ese caso, sigue el Protocolo de Servicio Suplementario (SS, Supplementary Service Protocol). La respuesta a una solicitud de localización en ASN.1 (Abstract Syntax Notation One) tiene una longitud de 133 bytes, asumiendo las Reglas Básicas de Codificación (BER, Basic Encoding Rules).

```
RESULT
 lcs-MOLRRes SEQUENCE {
   locationEstimate [0] IMPLICIT OCTET
       STRING ( SIZE (1 .. 20 ) )
     OPTIONAL,
   decipheringKeys [1] IMPLICIT OCTET
       STRING ( SIZE (15 ) ) OPTIONAL
     ,
   ... ,
   add-LocationEstimate [2] IMPLICIT
     OCTET STRING ( SIZE (1 .. 90 )
     ) OPTIONAL}
```

3. Análisis de escalabilidad

Consideramos que cada equipo de usuario genera una respuesta de localización cada vez que se une a una nueva microcelda (análisis de caso peor). Aproximamos la capacidad *Uu* necesaria a partir de la descripción dada en la sección previa.

En nuestro entorno de simulación, los usuarios de los equipo de usuario caminan de acuerdo al *modelo Manhattan* [12] en una región cubierta por microceldas. Todas las microceldas tienen el mismo radio. La tasa agregada *Iu* de tráfico de localización resulta de todos los cambios de microcelda de los equipos de usuario en el modelo. La Fig. 3 muestra la rejilla Manhattan.

Los usuarios caminan a lo largo de las líneas de la rejilla, y cambian su dirección en cada cruce con probabilidad 0.5. Además, cada cinco metros cambian su velocidad con probabilidad 0.2, de acuerdo a una distribución Normal, con media 3 Km/h y desviación estándar 0.3 Km/h.

A continuación, se estudia el escenario microcelular, y se compara con el caso en el que se despliega una RAL.

3.1. Configuración de la simulación del caso microcelular

Los valores específicos elegidos para este caso son:

- *Número de equipos de usuario*: 10-100.
- *Radio de las microceldas*: 100-500 metros.
- *Tamaño de la región*: 5000 × 5000 metros.

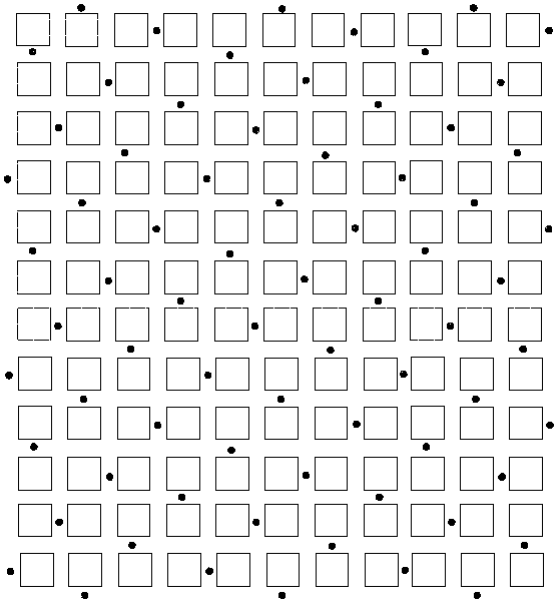


Figura 3: Rejilla Manhattan

- *Tiempo de simulación*: 100 segundos.

La Fig. 4 muestra el correspondiente tráfico de localización en la interfaz *Iu*.

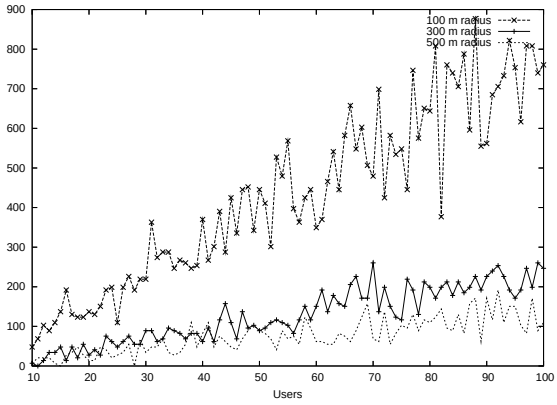


Figura 4: Escenario microcelular: tráfico de localización *Iu* (bps)

Como se podría esperar, la pendiente del tráfico de localización es inversamente proporcional al radio de las microceldas. Sin embargo, incluso para microceldas pequeñas (de radio 100 m) la tasa de tráfico de localización no pasa de 900 bps para 100 equipos de usuario. Nótese que la desviación estándar decrece con el tamaño de celda, dado que los equipos de usuario permanecen dentro de cada microcelda durante un tiempo mayor cuanto más grandes sean éstas.

Dado que la tasa de tráfico de localización es extremadamente baja, se podría tener la tentación de desplegar microceldas aún más pequeñas, pero, en ese caso, el incremento del tráfico *total* de señalización de (localización+llamada) puede llegar a ser inaceptable, debido a los traspasos de llamadas (handover). Llegados a este punto, decidimos introducir las picoceldas RAL, que son transparentes a las llamadas de usuario.

3.2. Configuración de la simulación de RAL

En el caso de RAL, los valores de la simulación son los siguientes:

- *Número de equipos de usuario*: 10-100.
- *Radio de las microceldas*: 10-100 metros.
- *Tamaño de la región*: 500 × 500 metros.
- *Tiempo de simulación*: 100 segundos.

El tráfico de localización en la interfaz *Iu* se muestra en la Fig. 5.

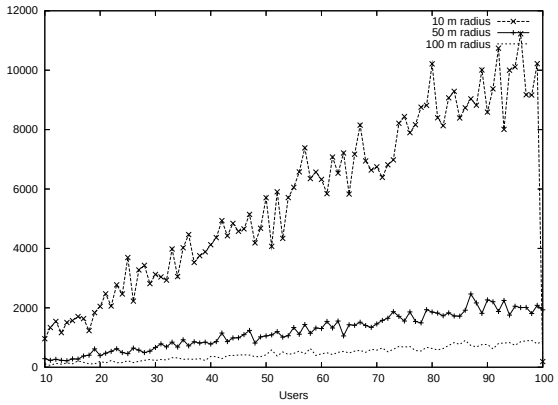


Figura 5: Escenario RAL: tráfico de localización *Iu* (bps)

En el caso peor (picocelda RAL de 10 m de radio), 100 usuarios en la rejilla Manhattan generan 12 Kbps de tráfico de localización. Esto es aproximadamente equivalente a 120 bps por equipo de usuario (para una tasa de datos de 384 Kbps, de acuerdo a los contratos actuales de Telefónica). Concluimos que el impacto de la integración de la RAL en la interfaz *Iu* es admisible en la práctica.

4. Integración RAL-UMTS

El *gestor RAL* es un subsistema del Nodo B controlado por un RNC. Cuando el RNC recibe una solicitud de localización en tiempo real, debe:

- Seleccionar el método de localización.
- Medir la posición del equipo de usuario periódicamente.

El RNC debe determinar si el Nodo B que sirve al equipo de usuario objetivo tiene una RAL. EL RNC empieza un proceso de averiguación usando una extensión del protocolo de la Parte de Aplicación de Nodo B (NBAP, Node B Application

Part) [13]. Si la respuesta del Nodo B es negativa, el RNC sigue un procedimiento de localización UMTS estándar. En caso contrario, el RNC sigue la pista del equipo de usuario con ayuda de la RAL. Resumiendo, el procedimiento que proponemos es como sigue:

- El Núcleo de Red (CN, Core Network) envía una solicitud de localización al UTRAN.
- El RNC recibe la solicitud, y determina el Nodo B que sirve al equipo de usuario objetivo.
- El RNC envía un comando `RAL_AVAILABILITY` y espera una respuesta:
  1. Negativa: el RNC retorna el Cell-ID del equipo de usuario al CN.
  2. Afirmativa:
     - El RNC envía el comando `RAL_POSITION_REQUEST`.
     - La RAL procesa la solicitud y retorna la posición del equipo de usuario objetivo con el comando `RAL_POSITION_REPORT`, inmediatamente o siempre que el área de localización cambie.
- El RNC construye mensajes de respuesta para el CN, (`LCS_AREA_EVENT_REPORT`).

En consecuencia, un conjunto de mensajes especializados controla a la RAL. Es posible añadir los nuevos mensajes al protocolo NBAP, tal como se muestra en la Tabla 1.

Tabla 1: Mensajes NBAP para la integración RAL-UMTS

IE/Group Name	**IE Type and Reference**	**Semantics Description**
Procedure ID		
Procedure Code	INTEGER (0..255)	"0" = Audit "1" = Audit Required "2" = Block Resource ... ... "50" = Bearer Rearrangement "51" = Radio Link Activación "52" = Radio Link Parameter Update **"53" = RAL Availability** **"54" = RAL Position Request** **"55" = RAL Position Report**
Ddmode	ENUMERATED (TDD, FDD, Common, ...)	Common = common to FDD and TDD
Type of Message	ENUMERATED (Initiating Message, Successful Outcome, Unsuccessful Outcome, Outcome)	

5. Conclusiones

En este escrito, proponemos la integración de redes de localización en interiores de alta precisión dentro de la arquitectura de UMTS, como una nueva entidad llamada Red Auxiliar de Localización. De este modo, se consigue aumentar la precisión de localización de forma totalmente transparente para la red UMTS (exceptuando el RNC). De acuerdo con el análisis de escalabilidad de la sección 3, el tráfico de localización RAL no tiene un impacto significativo sobre la interfaz *Iu*. Finalmente, proponemos una expansión NBAP para la integración RAL-UMTS.

Agradecimientos

Este trabajo ha sido financiado por Telefónica Móviles mediante un proyecto PPT-UMTS 2002.

Referencias

[1] 3GPP. Functional stage 2 description of localization services in UMTS. Technical report 3GPP TS 23.171 v3.9.0, 2002.

[2] P. Bahl and V. Padmanabhan. "RADAR: an in-building RF-based user location and tracking system." In *Proc. IEEE Infocom'00*, 775-784, 2000.

[3] F.J. González-Castaño and J.J. García-Reinoso. "Bluetooth Location Networks." In *Proc. IEEE Globecom'02*, 2002.

[4] J. Werb and C. Lanzl. "A positioning system for finding things indoors." *IEEE Spectrum* 35(9), 71–78, 1998.

[5] N. B. Priyantha, A. Chakraborty and H. Balakrishnan. "The Cricket location-support system." In *Proc. of the Sixth Annual ACM Intl. Conf. on Mobile Computing and Networking*, 2000.

[6] A. Harter, A. Hopper, P. Steggles, A. Ward and P. Webster. "The anatomy of a context-aware application." In *Proc. of the 5th Annual*

ACM/IEEE Intl. Conf. on Mobile Computing and Networking, 59–68, 1999.

[7] Texas Instruments TIRIS. http://www.ti.com/tiris/default.htm.

[8] Location Interoperability Forum. Mobile Location Protocol 3.0.0. [Online]. Available: http://www.openmobilealliance.org/tech/affiliates/LicenseAgreement.asp?DocName=/lif/LIF-TS-101-v3.0.0.zip

[9] 3GPP. UTRAN Iu Interface RANAP signaling. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/25413.htm

[10] A.B. García, M. Álvarez-Campana, E. Vázquez and J. Berrocal. "Simulation of Quality of Service Mechanism in the UMTS Terrestrial Radio Access Network." In *Proc. of the Fourth IEEE Conference on Mobile and Wireless Communications Networks (MWCN 2002)*, 2002.

[11] Geoff Varrall and Roger Belcher. *3G Handset and Network Design*. Wiley and Sons, 2003.

[12] 3GPP. Selection procedures for the choice of radio transmission technologies of the UMTS. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/3003U.htm

[13] 3GPP. UTRAN Iub interface NBAP signalling. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/25433.htm

PDP: Un protocolo de descubrimiento de servicios para redes ad hoc

Celeste Campo, Carlos García-Rubio, Andrés Marín, Florina Almenárez
Departamento de Ingeniería Telemática. Universidad Carlos III de Madrid
Avda. Universidad, 30. 28911 - Leganés (Madrid)
E-mail: {celeste, cgr, amarin, florina}@it.uc3m.es

Abstract *In ad hoc networks, mobile devices communicate via wireless links without any fixed infrastructure. These devices must be able to discover and share services dynamically. In this paper, we propose a new service discovery protocol specifically designed for this kind of networks, the Pervasive Discovery Protocol. PDP is a fully distributed protocol that merges characteristics of both pull and push solutions. Devices maintain a cache of services previously announced, that is also used for the answers. PDP reduces power consumption of the most limited devices by prioritising the replies of the less limited ones, allowing the others to abort their answers. Simulation results and an implementation are presented.*

1. Introducción

En los últimos años, los avances en microelectrónica y comunicaciones inalámbricas han propiciado la aparición de lo que se denomina "sistemas de computación ubicua", pequeños dispositivos con capacidad de proceso y de comunicación. Los más "visibles" de estos dispositivos son los teléfonos móviles y PDAs (*Personal Digital Assistants*), pero hay muchos otros que nos rodean que pasan más inadvertidos, como los electrodomésticos, muchos de los cuales tienen un procesador incorporado. Actualmente estos dispositivos ofrecen cada uno un servicio específico al usuario, pero en el futuro, gracias a su capacidad de comunicación, podrán colaborar con otros de su entorno para ofrecer conjuntamente servicios más complejos.

Los protocolos de descubrimiento de servicios hacen más sencillo descubrir y acceder al amplio número de servicios disponibles en una red (p.ej. impresoras, servidores de correo, etc.). Se han propuesto varios para Internet, dentro del IETF (SLP, SSDP y DNS-SD) o ligados a una tecnología concreta de alto nivel (Jini, Salutation). Más recientemente se han propuesto algunos específicamente diseñados para redes ad hoc, que revisaremos posteriormente. Sin embargo, se siguen necesitando nuevas soluciones de descubrimiento de servicios para redes ad hoc inalámbricas, adaptadas a su naturaleza cambiante y heterogénea. En estas redes, formadas en gran parte por dispositivos limitados, es importante minimizar el número total de transmisiones para así reducir el consumo de baterías. También es necesario detectar lo antes posible cuándo un servicio está disponible o deja de estarlo, a medida que nuevos dispositivos se unen o abandonan la red. Finalmente, dado que las transmisiones inalámbricas llegan a todos los dispositivos dentro del área de cobertura, es conveniente explotar al máximo la cooperación entre dispositivos para conseguir resultados óptimos.

Este artículo aborda el problema de anunciar y descubrir servicios en redes ad hoc formadas por dispositivos limitados. Nos centramos en redes ad hoc basadas en IEEE 802.11, que son las más ampliamente usadas. En ellas, los nodos deben estar siempre activos para poder procesar en cualquier momento una petición de servicio, no se pueden usar modos de bajo consumo (*sleep*). Proponemos un nuevo protocolo de descubrimiento de servicios, *Pervasive Discovery Protocol* (PDP), definido para satisfacer los requisitos de estas redes. El artículo está organizado como sigue. En el apartado 2 revisamos las aproximaciones teóricas al problema del descubrimiento de servicios, y los protocolos ya existentes. En el apartado 3 enumeramos los nuevos requisitos y retos que imponen las redes ad hoc al descubrimiento de servicios. En el apartado 4 revisamos los protocolos de descubrimiento de servicios definidos específicamente para redes ad hoc. En el apartado 5 introducimos nuestra propuesta, el protocolo PDP. En el 6 presentamos resultados de simulación que comparan nuestro protocolo con las aproximaciones teóricas y muestran diversos aspectos de las prestaciones de PDP. En el apartado 7 presentamos la implementación de PDP en J2ME. Finalmente, en el 8, incluimos las conclusiones y trabajo futuro.

2. Soluciones para el descubrimiento de servicios

El descubrimiento de servicios no es un problema nuevo. Se han propuesto múltiples soluciones, con distinto grado de aceptación. En este aparta-

do revisamos las aproximaciones teóricas posibles y presentamos brevemente los protocolos existentes para redes de cable.

2.1. Aproximaciones

El descubrimiento de servicios aborda el problema en el que ciertos elementos de la red, los clientes, necesitan conocer los servicios ofrecidos por otros elementos, los servidores. Este problema puede abordarse siguiendo una aproximación distribuida o centralizada.

En la aproximación distribuida no se requiere ningún dispositivo adicional. El descubrimiento se puede hacer de dos maneras: (i) modo *push*, los servidores anuncian periódicamente sus servicios, los clientes escuchan estos anuncios y seleccionan aquellos que les interesan; (ii) modo *pull*, los clientes piden un servicio cuando lo necesitan, los servidores que ofrecen ese servicio contestan.

En la aproximación centralizada se introduce un nuevo elemento en la red, el directorio (que puede ser físicamente una única máquina o una jerarquía de máquinas). Los servidores registran sus servicios en el directorio; los clientes mandan a éste sus peticiones de servicio. De esta manera se minimizan las transmisiones.

Los protocolos que veremos en los siguientes apartados responden a una o más de estas aproximaciones.

2.2. Protocolos existentes

El descubrimiento de servicios se trató por primera vez en Internet en 1993, en el grupo SRVLOC del IETF. El principal resultado de este grupo fue la definición del *Service Location Protocol* (SLP). SLP versión 2 es actualmente un estándar Internet (RFC 2608). SLP define tres tipos de "agentes": el Agente de Usuario (UA), que descubre servicios, el Agente de Servicio (SA), que anuncia servicios, y el Agente Directorio (DA), que almacena información sobre los servicios anunciados. SLP soporta dos modos de funcionamiento: si hay un DA en la red sigue la aproximación centralizada, si no lo hay sigue la distribuida con modo *pull*.

Dentro de SRVLOC se propuso otro protocolo, el *Simple Service Discovery Protocol* (SSDP) [1], que pese a que no pasó del estado de draft, es el que se usa en la arquitectura UPnP de Microsoft. SSDP define un protocolo mínimo de descubrimiento basado en multicast. La versión del IETF soportaba funcionamiento *pull* y *push*, en la versión UPnP se incluyó el basado en directorio.

También en 1999, UC Berkeley propuso el *Secure Service Discovery Service* (SSDS) [2], que es un protocolo tolerante a fallos y escalable enfocado a redes de área extensa. SSDS soporta los modos *pull*, *push* y centralizado.

Otras propuestas están ligadas a una tecnología de alto nivel de desarrollo de aplicaciones, como Jini [3] y Salutation [4]. Ambas se basan en la existencia de uno o más directorios.

3. Retos en redes ad hoc

Los protocolos de descubrimiento que hemos presentado en el apartado anterior se diseñaron para usarse en Internet (excepto quizás SSDP), con redes de cable y ordenadores fijos. En las redes ad hoc los dispositivos se comunican de forma inalámbrica, sin necesidad de infraestructura fija; muchos de ellos son móviles y tienen distintas características en cuanto a batería, CPU y memoria; las redes tienen una topología dinámica y requieren una gestión descentralizada. Existe un gran consenso en la literatura en que las soluciones de descubrimiento de servicios existentes no abordan bien estos problemas [5].

Recientemente se han propuesto algunos protocolos de descubrimiento específicos para redes ad hoc. Los revisaremos en el siguiente apartado, antes presentaremos aquí los retos más importantes que se debe abordar:

Minimizar las transmisiones en la red. Los dispositivos en estas redes frecuentemente se alimentan con baterías, cuyo consumo debe limitarse al máximo. En una red 802.11 ad hoc todos los nodos deben estar siempre en modo activo para atender posibles peticiones. Una de las principales fuentes de consumo es la transmisión [6] (una tarjeta WLAN PCMCIA convencional consume 1.5 W en modo transmisión). Esto implica que debe diseñarse el protocolo de manera que se minimice el número de transmisiones, especialmente de los dispositivos más limitados. Una posible solución es introducir una pequeña caché en cada dispositivo; sin embargo en redes muy dinámicas, como las ad hoc, las cachés deben usarse con precaución porque pueden quedar desactualizadas rápidamente.

No depender de infraestructura. Los protocolos que utilizan directorio no son adecuados en redes ad hoc ya que éstas se forman allá donde dos o más dispositivos coinciden, no siempre se va a disponer de una infraestructura (un directorio) en la que apoyarse. Una posible solución es elegir dinámicamente un directorio entre los dispositivos de la red, pero es costoso en tiempo y en tráfico. Por tanto en redes ad hoc los protocolos de descubrimiento deben seguir la aproximación distribuida, pero intentando ser eficientes tanto en entornos muy cambiantes como en estáticos.

Aprovechar la capacidad de difusión. En las redes ad hoc el canal físico es compartido por todos los dispositivos; el coste de mandar un paquete a un dispositivo (*unicast*) o a todos los dispositivos en rango (*broadcast*) es el mismo. En redes ad hoc de múltiple salto (MANETs) se han propuesto protocolos de encaminamiento multicast [7] que soportan la difusión de modo eficiente. SLP, SSDP y gran parte de los protocolos para redes ad

hoc que veremos después explotan esta característica.

Maximizar la cooperación entre dispositivos. La cooperación es un concepto clave en computación ubicua, que permite a dispositivos limitados llevar a cabo conjuntamente tareas complejas. Esto aplicado al descubrimiento de servicios implica que terceros dispositivos, no sólo el que busca y el que ofrece un servicio, pueden cooperar en la búsqueda o beneficiarse de ella.

Adaptarse a entornos muy cambiantes. Las redes ad hoc son mucho más dinámicas y cambiantes que Internet, y la información en caché queda obsoleta rápidamente. Algunos protocolos de descubrimiento existentes ya incluyen mecanismos para garantizar la consistencia de la caché, como el tiempo de vida de los servicios en SLP. Sin embargo, estos mecanismos no son suficientes para redes ad hoc, y se hace necesario proponer nuevas soluciones.

Tener en cuenta las diferentes necesidades de las aplicaciones. Finalmente podemos distinguir dos tipos de aplicaciones que desean buscar un servicio: aquellas que quieren encontrar un dispositivo cualquiera que lo ofrezca, da igual cuál, y aquellas que necesitan conocer todos los dispositivos que ofrecen el servicio en la red. Los protocolos de descubrimiento hasta la fecha no han distinguido ambos tipos de consulta, lo que supone un derroche de ancho de banda.

4. Protocolos de descubrimiento ad hoc

Los protocolos de descubrimiento preexistentes que hemos revisado en el apartado 2 se definieron para trabajar en redes de cable, aunque algunos de ellos han sido adaptados a redes inalámbricas con dispositivos limitados, por ejemplo Salutation Lite o JiniME.

En los últimos años han aparecido nuevas propuestas que abordan el problema del descubrimiento de servicios en redes ad hoc. Algunas propuestas están ligadas a una tecnología de subred, como *Service Discovery Protocol* de Bluetooth e *Information Access Service* de IrDA. Otras abordan el problema del descubrimiento de forma conjunta al del encaminamiento, como GSD, HSID y el trabajo Koodli y Perkins. Otras trabajan al nivel de aplicación de la pila de protocolos, éstas son las que revisaremos con más detenimiento más adelante. Otras están ligadas a una tecnología de alto nivel de desarrollo de aplicaciones distribuidas, como el mecanismo de descubrimiento en JXTA Finalmente otras propuestas forman parte de una arquitectura de computación ubicua más global, como one.world e INS de Oxygen.

Desde nuestro punto de vista, el descubrimiento de servicios debe realizarse al nivel de aplicación, por las siguientes razones. Una solución vinculada a una tecnología de subred o de desarrollo de aplicaciones es válida únicamente en redes homogéneas, mientras que las redes ad hoc son heteogéneas por naturaleza. En cuanto a las soluciones conjuntas de descubrimiento y encaminamiento, nosotros creemos que el problema del encaminamiento unicast y multicast en MANETs es muy importante y debe ser abordado de forma general y no ligado a una aplicación concreta, por importante que ésta sea. Defendemos, por tanto, que el problema del descubrimiento de servicios debe solucionarse al nivel de aplicación. Ésta es también la aproximación que ha sido adoptada por el IETF y por otros grupos de investigación, que presentamos ahora más detalladamente.

El algoritmo DEAPspace [8] propone una solución *push* en la que los dispositivos mantienen una lista de servicios conocidos, lo que llama su "vista del mundo", que difunde periódicamente a sus vecinos. La evaluación de prestaciones realizada sobre DEAPspace demuestra que el ancho de banda que consume es similar al modo *push*, y por tanto depende de la frecuencia de anuncios. Sin embargo, el tiempo consumido para descubrir los servicios disponibles es menor ya que la información que se propaga en cada anuncio es mayor.

Rendezvous [9] permite el descubrimiento automático de ordenadores, dispositivos y servicios en redes IP. Rendezvous, también conocido como *zero-configuration networking*, permite a dos dispositivos IP comunicarse sin necesidad de configurar previamente sus direcciones ni tener un servidor DNS. La parte de descubrimiento se basa en DNS. Para poder funcionar en redes sin infraestructura, RendezVous define un DNS distribuido denominado Multicast DNS (draft de IETF), que se comporta como un *pull* pero con respuestas multicast. Hay otra propuesta similar de DNS distribuido, LLMNR, que se está desarrollando dentro del grupo DNSEXT del IETF.

Konark [10] es un *middleware* diseñado para descubrir y proporcionar servicios en redes ad hoc multisalto. Con respecto al descubrimiento, Konark soporta tanto el modo *push* como el *pull*, los servidores pueden anunciar servicios y los clientes pedirlos. Todos los dispositivos tienen una caché con los servicios locales del dispositivo y los remotos previamente anunciados. En el modo *pull*, los dispositivos pueden responder con la información de su caché, aunque sea un servicio no local. Igualmente los anuncios pueden incluir servicios locales y remotos. Konark ha definido recientemente un nuevo mecanismo de descubrimiento denominado *Konark Service Gossip Protocol*, basado en una ronda de intercambio de mensajes.

4.1. Características y carencias

Los protocolos que acabamos de revisar presentan algunas características que nuestro protocolo

va a reaprovechar, pero también algunas carencias que nuestra propuesta pretende superar.

Entre las características, podemos señalar las siguientes: todos ellos (i) no dependen de infraestructura, siguen la aproximación distribuida; (ii) hacen uso del soporte broadcast / multicast subyacente; (iii) incluyen un temporizador aleatorio para evitar colisiones cuando un dispositivo contesta una petición; (iv) usan cachés en los dispositivos para minimizar el número de tranmisiones; (v) anuncian los servicios con un tiempo asociado, que es el que la entrada podrá permanecer en caché; (vi) todos, excepto Rendezvous, se basan en el modo *push*, aunque algunos soportan además el *pull* para cuando la exactitud es importante (la información de la caché de *push* puede estar caducada); (vii) finalmente, todos excepto Rendezvous, contestan las peticiones de servicios no sólo con los servicios que ofrece el dispositivo localmente, sino también con los de la caché.

Las carencias observadas, y que nuestro protocolo aborda, son: (i) no tienen en cuenta las diferentes características de los dispositivos (batería, memoria); (ii) todos, excepto Rendezvous, al basarse en *push*, transmiten anuncios incluso cuando no hay nadie más en la red; (iii) soportan el modo *pull*, pero no está bien integrado con el *push*, las aplicaciones deben escoger entre usar uno u otro. (iv) no tienen en cuenta las diferentes necesidades de las aplicaciones; (v) finalmente, como una parte importante de los dispositivos tienen una gran movilidad, se necesita mecanismos para garantizar la consistencia de las cachés. Rendezvous soporta uno de estos mecanismos. Cuando un dispositivo abandona la red, puede transmitir un anuncio de todos sus servicios con TTL = 0 para que así los que lo escuchen borren los anuncios previos de sus cachés. Nosotros incluiremos también este tipo de mecanismos en nuestro protocolo.

5. Pervasive Discovery Protocol

En este apartado presentamos la definición e implementación de un nuevo protocolo de descubrimiento de servicios, el Pervasive Discovery Protocol (PDP), que ha sido especialmente diseñado para funcionar en redes ad hoc. Una versión preliminar del protocolo la presentamos en [11].

El protocolo que proponemos no necesita un directorio central. Uno de los objetivos principales de PDP es minimizar el gasto de batería en todos los dispositivos, especialmente en los más limitados. Esto significa que el número de transmisiones necesarias para descubrir servicios debería reducirse el máximo posible. Un dispositivo anunciará sus servicios sólo cuando otro dispositivo solicite el servicio. Los anuncios de servicios se mandan por difusión a todos los dispositivos en rango, de manera que todos ellos conocerán el servicio aunque no hayan preguntado por él. Los dispositivos poseen una caché donde se almacenan los anuncios de servicios. Los anuncios incluyen no sólo los servicios ofrecidos por el propio dispositivo que lo envía, sino todos los servicios que conoce (ofrecidos por otros) del tipo solicitado. PDP tiene en cuenta la heterogeneidad de los dispositivos, reduciendo el consumo de los más limitados, esto se consigue priorizando la respuesta de los menos limitados y permitiendo a los otros abortar su respuesta si no tienen nada nuevo que decir.

Ahora definiremos PDP más formalmente.

5.1. Escenario de aplicación

Supongamos que tenemos una red ad hoc con `D` dispositivos, cada uno ofrece `S` servicios, almacenados en un array, `Local`, y espera permanecer en la red durante `T` segundos. Este tiempo `T` se denomina *tiempo de disponibilidad* y ha sido previamente configurado en el dispositivo, dependiendo de sus características de movilidad. Este parámetro, con éste u otro nombre (por ejemplo TTL) es habitual en todos los protocolos de descubrimiento de servicios. En nuestro protocolo se asigna un tiempo de disponibilidad menor a los dispositivos más limitados (con menos memoria y que funcionan con batería) pues suelen ser los más móviles.

Cada dispositivo tiene un Agente de Usuario PDP (PDP_UA) y un Agente de Servicio PDP (PDP_SA). El PDP_UA en un proceso que se encarga de hacer la búsqueda de servicios en la red. El PDP_SA es un proceso que anuncia los servicios que ofrece el dispositivo. El PDP_SA siempre incluye el tiempo de disponibilidad `T` del dispositivo que lo anuncia.

Cada dispositivo tiene una `Cache` que contiene la lista de servicios que han sido escuchados hasta el momento en la red. Cada elemento de la caché tiene dos campos: la *descripción del servicio* y el *tiempo de expiración del servicio*. Este último es el tiempo estimado que resta al servicio estar disponible en la red; se calcula como el mínimo de dos valores: el tiempo de disponibilidad del dispositivo local, y el tiempo de vida anunciado del servicio. Las entradas de la caché se borran cuando vence su tiempo de expiración.

5.2. Descripción del protocolo

PDP tiene dos mensajes obligatorios: `PDP_Service_Request`, que se usa para enviar peticiones de servicios, y `PDP_Service_Reply`, para anunciar servicios. PDP define además un mensaje opcional, `PDP_Service_Deregister`, que se usa para informar de que un servicio dejará de estar disponible en el futuro.

Ahora explicaremos en detalle cómo usan el PDP_UA y el PDP_SA estas primitivas.

5.2.1. Agente de Usuario PDP

Cuando una aplicación o el usuario final de un dispositivo necesita un servicio de un cierto tipo, hace una llamada a su PDP_UA. Para soportar eficientemente distintos tipos de aplicaciones, en PDP hemos definido dos tipos de consultas:

- **una petición–una respuesta** (1/1): la aplicación está interesada en el servicio, no en qué dispositivo lo ofrece.
- **una petición–multiples respuestas** (1/n): la aplicación quiere descubrir todos los dispositivos de la red que ofrecen un cierto servicio. Para esta petición hemos incluido un tipo especial de servicio, denominado `ALL`, para permitir a la aplicación descubrir todos los servicios de todos los tipos en la red.

Ambos tipos de consulta usan el mismo mensaje, `PDP_Service_Request`, un bit en la cabecera indica si se trata de 1/1 o 1/n.

En **una petición–una respuesta**, el PDP_UA busca el `service_type` especificado en la lista de servicios locales y en su caché. Si lo encuentra, el PDP_UA devuelve a la aplicación la descripción de servicio correspondiente, sin transmitir a la red. Si no lo encuentra, el PDP_UA difunde un `PDP_Service_Request` de ese servicio y espera `CONFIG_WAIT_RPLY` segundos a que lleguen respuestas. Si no llega ninguna, el PDP_UA responde a la aplicación que el servicio no está disponible en la red. Si llega alguna, el PDP_UA devuelve a la aplicación la descripción del servicio recibida.

En **una petición–multiples respuestas**, el PDP_UA construye una lista de servicios conocidos del tipo buscado (todos los servicios si el tipo es `ALL`), esto es, los locales y los que tenía almacenados en la caché. A continuación envía un `PDP_Service_Request` incluyendo dicha lista. Espera `CONFIG_WAIT_RPLY` segundos a que lleguen respuestas, y cuando pasa este tiempo entrega a la aplicación la lista de servicios conocidos: la inicial más aquellos que se hayan recibido.

Los PDP_UA de todos los dispositivos escuchan continuamente la red (tanto las peticiones como las respuestas) y actualizan la caché del dispositivo con los servicios que escuchan. La caché tiene un tamaño limitado; cuando un PDP_UA escucha un anuncio pero la caché está llena, borra el servicio al que resta menos para expirar y almacena el servicio anunciado.

5.2.2. Agente de Servicio PDP

El PDP_SA anuncia los servicios ofrecidos por el dispositivo. Su función es procesar los mensajes `PDP_Service_Request` y generar el correspondiente `PDP_Service_Reply`, si es necesario.

Para minimizar el número de transmisiones, el PDP_SA tiene en cuenta el tipo de la consulta. Cuando recibe un `PDP_Service_Request 1/1`, comprueba si el servicio solicitado es uno de sus servicios locales. En ese caso, planifica la transmisión de un `PDP_Service_Reply` dentro de un tiempo aleatorio, inversamente proporcional al tiempo de disponibilidad del dispositivo. Durante este tiempo, si escucha otra respuesta a la misma petición, aborta su `PDP_Service_Reply` ya que el PDP_UA remoto lo descartaría. Si el temporizador expira sin que nadie más haya respondido, la transmite. El algoritmo da a los dispositivos más estáticos, con mayor tiempo estimado de disponibilidad, más oportunidad de responder primero.

Cuando un PDP_SA recibe un `PDP_Service_Request 1/n`, comprueba si el servicio solicitado en uno de sus servicios locales, o si está en su caché. Si es así, genera un tiempo de espera aleatorio, inversamente proporcional al tiempo de disponibilidad del dispositivo y al número de servicios conocidos. Durante este tiempo, el PDP_SA escucha en la red otros posibles `PDP_Service_Reply` de otros dispositivos a la misma petición. Cuando el temporizador expira, si el PDP_SA conoce algún servicio que todavía no ha sido anunciado, envía su `PDP_Service_Reply`. De esta manera, cuanto más tiempo de disponibilidad tiene el dispositivo y mayor es su caché, mayor es la probabilidad de que conteste primero. Estamos suponiendo que este dispositivo es el que tendrá la visión más completa de los servicios que hay en la red ad hoc.

Puede ocurrir que cuando apagamos un dispositivo, o nos vamos a cambiar de red, el dispositivo tenga tiempo suficiente para enviar un último mensaje. Si es así, el PDP_SA del dispositivo puede transmitir el mensaje opcional `PDP_Service_Deregister` con todos sus servicios. Cuando el resto de dispositivos escuchen este mensaje, borrarán de sus cachés los servicios del dispositivo. Este mensaje puede enviarlo también cualquier dispositivo cuando intenta acceder a un servicio que tenía en su caché, y descubre que dicho servicio ya no está disponible en la red.

5.3. Otras consideraciones

Todos los mensajes PDP se transmiten por multicast, usando el protocolo UDP. Si el mensaje completo no cabe en un datagrama, se fragmenta en varios mensajes que se envían de forma independiente. El puerto en el que escucha PDP es el 3000 (no está reservado por IANA). Este es el puerto destino de todos los mensajes PDP. Para la transmisión se usa la dirección multicast 239.255.255.253 (RFC2365). El TTL multicast por defecto es 255. En redes aisladas, por ejemplo en redes ad hoc de un sólo salto, puede usarse broadcast en lugar de multicast. El valor por defecto del temporizador `CONFIG_WAIT_RPLY` es 3 segundos. El tiempo aleatorio que un PDP_SA espera antes de enviar una respuesta se genera usando la ecuación 1, donde $U(x, y)$ es una distribución uniforme

entre x e y, T_D es el tiempo de disponibilidad del dispositivo y Service_Entries_Number es el número de servicios conocidos, que se incluirá en la respuesta.

$$U(0, \quad \texttt{CONFIG\_WAIT\_RPLY} * \frac{3600}{3600 + T_D * \text{Service\_Entries\_Number}}) \quad (1)$$

La definición completa del protocolo PDP, incluyendo formato de mensajes, encapsulación en UDP, puertos de transporte, direcciones multicast, etc., está disponible en [12]. Para la descripción de los servicios no definimos un nuevo formato sino que reutilizamos el definido por el grupo SRVLOC del IETF como parte de su protocolo SLP (RFC2609).

6. Evaluación de prestaciones

En este apartado presentamos un estudio de prestaciones de PDP en una red ad hoc. Comparamos nuestro protocolo con las aproximaciones teóricas: *push*, *pull* y directorio. Este estudio lo hemos llevado a cabo mediante simulación.

6.1. Entorno de simulacion

Hemos realizado una implementación del protocolo PDP en ns-2 y un simulador específico de PDP en MODSIM [13]. En éste último hemos realizado simulaciones de horizonte infinito, empleando el método de medias por bloques (*batch-means*) con un nivel de confianza del 90 % y un intervalo de confianza del 10 %.

Simulamos una red 802.11 ad hoc multisalto con soporte de routing multicast, en la que aparecen dispositivos nuevos de forma aleatoria, solicitan y ofrecen servicios de forma aleatoria y abandonan la red después de un tiempo aleatorio. El número de dispositivos que existen en la red, varía con el tiempo, pero su media permanece estacionaria. Los tiempos aleatorios siguen una distribución exponencial, mientras que la asignación de servicios ofrecidos a dispositivos sigue una distribución uniforme. Por simplicidad, consideramos que cada dispositivo ofrece un único servicio.

Los parámetros de las simulaciones realizadas son: número medio de dispositivos, tiempo medio que permanece el dispositivo en la red, tamaño de la caché, tiempo medio entre búsqueda de servicios y el número de tipos de servicios en la red. Los resultados de interés son: número de mensajes transmitidos (normalizado al número de búsqueda de servicios), tasa de descubrimiento de servicios (porcentaje de servicios descubiertos respecto al real) y tasa errores (servicios falsos descubiertos, que en realidad no están disponibles en la red).

6.2. Resultados de simulación

Compararemos en primer lugar PDP con *pull* y *push* en una red ad hoc. La figura 1 muestra el número de mensajes transmitidos, la tasa de descubrimiento de servicios y la tasa de errores en un escenario con una media de 20 dispositivos, un tiempo medio de vida de los dispositivos entre 600 y 19200 segundos, un tamaño de caché de 100 entradas, 5 tipos de servicio diferentes y en el que cada dispositivo busca un servicio en media cada 60 segundos. El número de mensajes de PDP está muy por debajo de los obtenidos en *pull*, con la misma tasa de descubrimiento y de errores. En la figura vemos también las prestaciones de dos protocolos de descubrimiento *push*, uno con periodo entre anuncios de 60 segundos, que iguala el número de mensajes enviados por PDP pero con peor tasa de descubrimiento y de errores, y otro con periodo 12 segundos, que iguala la tasa de descubrimiento y de errores de PDP pero que transmite cinco veces más mensajes.

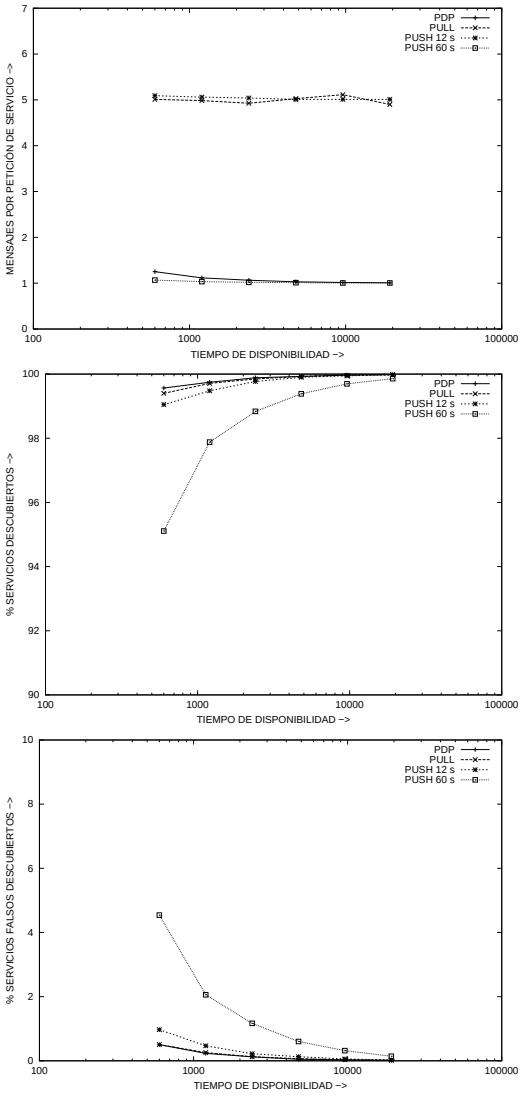


Figura 1: Comparación de PDP con otros protocolos.

Ahora estudiaremos el impacto del número de dispositivos en la red y del tamaño de la caché en las prestaciones de PDP. PDP con cache 0 es equivalente a un modo *pull*. La figura 2 muestra que si el tamaño de caché es suficientemente grande, el número de mensajes transmitidos permanece constante. Para tamaños de caché pequeños, cuando el número de dispositivos iguala el tamaño de la caché, el número de mensajes comienza a aumentar linealmente. Para caché 0 (modo *pull*) el incremento es siempre lineal.

Ahora demostraremos cómo PDP tiene en cuenta la herogeneidad de los dispositivos, consiguiendo una reducción de tráfico transmitido y de consumo energético de los dispositivos más limitados. La figura 3 muestra la tasa de respuestas enviadas por cada clase de dispositivo, dependiendo de su tiempo de disponibilidad. Hemos considerado un escenario con una media de 40 dispositivos, con 5 valores diferentes de tiempos de disponibilidad: 500, 2500, 4500, 6500 y 9500 segundos, con (en media) un 20 % de dispositivos de cada tipo. El resto de los parámetros son los mismos que en las simulaciones anteriores, excepto que el tamaño de caché de los dispositivos con disponibilidad 500 y 2500 es 10 servicios, de los de 4500 y 6000 es 40 servicios, y de los de 9500 es 100 servicios.

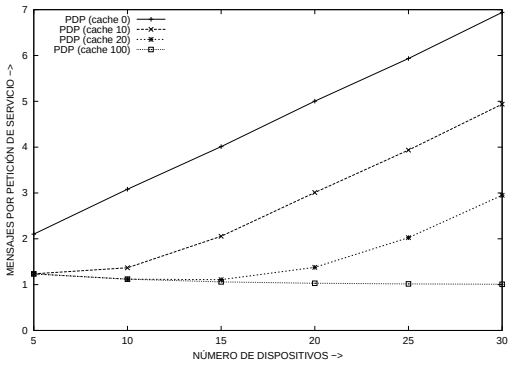


Figura 2: Tasa de respuestas por búsqueda para diferente númeror de dispositivos.

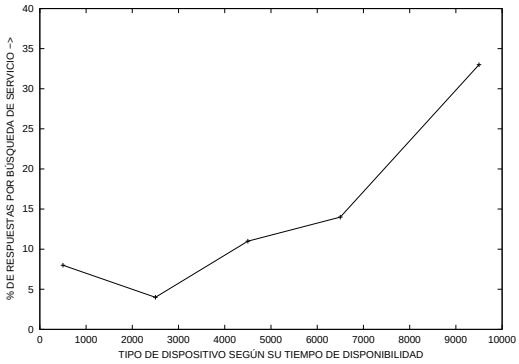


Figura 3: Tasa de respuestas por búsqueda con dispositivos de diferentes clases.

En esta figura vemos que los dispositivos con mayor tiempo de disponibilidad contestan más peticiones, lo que permite ahorrar batería a los más limitados. Hay que señalar que la suma de los porcentajes de respuesta a peticiones para todos los dispositivos, es del 70 % ya que en PDP algunas peticiones no generan respuesta (cuando todos los servicios conocidos estaban incluidos en la propia petición). Teniendo esto en cuenta, vemos que los dispositivos con disponibilidad de 9500 segundos contestan casi la mitad de las peticiones. Con cualquiera de los otros protocolos de descubrimiento que hemos mencionado, todos los dispositivos habrían contestado con igual probabilidad, 20 %. Esto significa que en PDP los dispositivos fijos y menos limitados contestan a más peticiones, como nos habíamos marcado en los objetivos de nuestro protocolo. La figura 3 también muestra que los dispositivos con tiempo de disponibilidad pequeño (en nuestro caso, 500 segundos) contestan más peticiones que los de tiempos intermedios. Ésto se debe a que los dispositivos de tiempo de disponibilidad pequeño (500 segundos) son muy móviles, contínuamente cambian de red, y cada vez que llegan a un nuevo entorno tienen que contestar para anunciar sus servicios y darse a conocer.

En la figura 4 mostramos una estimación del consumo energético de cada dispositivo en el mismo escenario de antes, en función de su tiempo de disponibilidad. Para obtener este valor, hemos utilizado los datos de consumo de una tarjeta de red inalámbrica publicados en [6]. Los resultados están normalizados al coste de transmisión de un mensaje. Vemos que en PDP los dispositivos menos limitados consumen 25 $\mu W.sec$ más que los más limitados. El consumo varía entre 77 y 102 $\mu W.sec$; de éstos, 76 $\mu W.sec$ corresponden a recepciones y el resto a transmisiones. En un modo *pull*, todos los dispositivos consumen lo mismo: 241 $\mu W.sec$, mientras que en un *push* con anuncios cada 60 segundos el consumo es de 861 $\mu W.sec$.

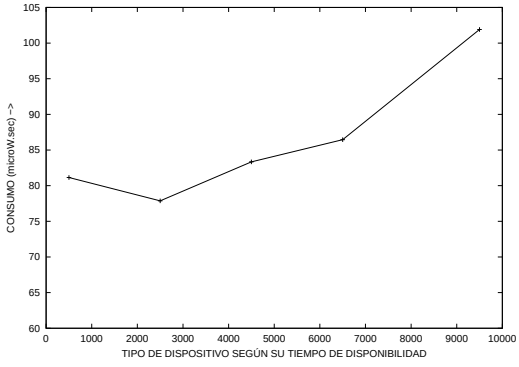


Figura 4: Consumo energético para distintos tipos de positivos.

7. Aspectos implementación

Hemos realizado una implementación de PDP tanto en *Java 2 Standard Edition* (J2SE), para ordenadores de sobremesa y portátiles, como en *Java 2 Micro Edition* (J2ME) para dispositivos limitados. Centrándonos en la implementación J2ME,

utilizamos el perfil *Personal Profile* de la configuración *Connected Device Configuration* (CDC). Elegimos este perfil y configuración porque soporta transmisiones multicast. Para el desarrollo hemos usado el *IBM WebSphere Studio Device Developer*. Se ejecuta en todos los dispositivos soportados por la máquina virtual *WebSphere Micro Environment* (WME), también conocida como J9. El tamaño del `jar` que incluye tanto el PDP_UA como el PDP_SA es de 19 KB. Hemos probado la implementación en ordenadores de sobremesa, portátiles y PDAs Pocket PC, todos ellos con interfaces IEEE 802.11b. Se puede encontrar más información sobre la implementación y descargarla bajo licencia LGPL en [12].

8. Conclusiones

En las redes ad hoc, los dispositivos deben ser capaces de descubrir y compartir dinámicamente servicios. Los protocolos de descubrimiento de servicios existentes no se diseñaron para entornos tan dinámicos. En este artículo proponermos un nuevo protocolo de descubrimiento de servicios, PDP, especialmente adecuado para redes ad hoc. PDP es un protocolo distribuido que fusiona características de *pull* y de *push*, y que explota la cooperación entre dispositivos para reducir al mínimo el tráfico y el consumo de batería, sin penalizar la tasa de servicios descubiertos.

Una posible crítica a PDP sería su vulnerabilidad a ataques de seguridad. Por ejemplo, un dispositivo malicioso podría anunciar servicios falsos o enviar falsos `PDP_Service_Deregister` de servicios de terceros. Como en otros protocolos de descubrimiento (SLP, Multicast DNS o LLNMR), PDP asume que los dispositivos en la red ad hoc son amistosos y cooperan por el bien común. La protección frente a dispositivos maliciosos en estos entornos no es un problema únicamente del descubrimiento de servicios, sino de todos los protocolos a todos los niveles (por ejemplo, ARP). Para resolver este problema se necesita un mecanismo más genérico que permita determinar qué dispositivos son de confianza y cuáles no. Hemos propuesto uno de estos mecanismos, basado en un modelo de confianza para redes ad hoc, en [14].

Como trabajo futuro, estamos trabajando en su implementación en dispositivos que no soporten la Java Virtual Machine, como cámaras IP; en el aprendizaje dinámico y autoconfiguración del tiempo de disponibilidad de los dispositivos (actualmente en PDP, como en los otros protocolos de descubrimiento, este valor se configura manualmente); y en dar soluciones a los problemas de seguridad inherentes a las redes ad hoc.

Agradecimientos

Este trabajo ha sido parcialmente soportado por los proyectos EVERYWARE (TIC2003-08995-C02-01), UBISEC (IST-2002-506926) y EASY WIRELESS (ITEA ip03008). Agradecemos a los revisores anónimos sus comentarios.

Referencias

[1] Y.Y. Goland, T. Cai, P. Leach, and Y. Gu. Simple Service Discovery Protocol/1.0. Internet-Draft (work in progress), April 1999. draft-cai-ssdp-v1-03.txt.

[2] S.E. Czerwinski, B.Y. Zhao, T.D. Hodes, A D. Joseph, and R.H. Katz. An Architecture for a Secure Service Discovery Service. In *Proc. Mobicom'99*, 1999.

[3] http://www.sun.com/jini/whitepapers/architecture.pdf.

[4] http://www.salutation.org.

[5] S. Helal. Standards for Service Discovery and Delivery. *IEEE Pervasive Computing*, pages 95–100, jul/sep 2002.

[6] L.M. Feeney and M.Ñilsson. Investigating the Energy Consumption of a Wireless Network Interface in an Ad Hoc Networking Environment. In *IEEE INFOCOM*, 2001.

[7] C. Morais, H. Gossain, and D.P. Agrawal. Multicast over wireless mobile ad hoc networks: present and future directions. *IEEE Network*, 17(1):52–59, Jan.-Feb. 2003.

[8] M.Ñidd. Service Discovery in DEAPspace. *IEEE Personal Communications*, August 2001.

[9] http://developer.apple.com/macosx/rendezvous/.

[10] S. Helal, N. Desai, and V. Verma. Konark-A Service Discovery and Delivery Protocol for Ad-hoc Networks. In *Third IEEE Conference on Wireless Communication Networks (WCNC)*, New Orleans, March 2003.

[11] C. Campo, M. Munoz, J.C. Perea, A. Marin, and Garcia-Rubio. PDP and GSDL: A New Service Discovery Middleware to Support Spontaneous Interactions in Pervasive Systems. In *IEEE PerWare. PerCom 2005 Workshops*, Kauai, Hawaii, March 2005.

[12] http://www.it.uc3m.es/celeste/tesis/.

[13] Brian Wood and Kerim Tumay. Modsim iii and caci's applications. In *WSC '99*, pages 234–238, New York, NY, USA, 1999. ACM Press.

[14] F. Almenarez and C. Campo. SPDP: A Secure Service Discovery Protocol for Ad-hoc networks. In *EUNICE 2003*, pages 213–218, Hungary, Budapest - Balatonfüred, September 2003.

Reconfiguración de espacios inteligentes mediante la integración de tarjeta inteligentes

Juan Jesús Sánchez Sánchez, José Alberto Vigo Segura,
Natividad Martínez Madrid, Ralf Seepold
Departamento de Ingeniería Telemática
Universidad Carlos III de Madrid
Av. de la Universidad, 30
28911 Leganés, Madrid, Spain
{jjsanchez,jvigo,nati,ralf}@it.uc3m.es

Abstract. *Current building market trends show that traditional building equipment is not sufficient to satisfy the needs of its customers. A fast-moving sector where these kinds of innovative equipment are being introduced is in commercial and public buildings like offices, meetings rooms, conferences, airports, etc. Here, rooms and spaces are required to be able to modify its configuration on demand. Automation networks can be used to support this capability. Commonly, reconfiguration of the spaces is provided through a remote control. This kind of solution has some restrictions like preferences management, security, energy consumption and dimensions. In this article, an alternative solution based on smart cards is proposed that allows users to store their profiles in the card. This model integrates smart cards into a control network structure in order to setup, store and reconfigure smart spaces. As a result, users can share and transfer profiles in different environments without accessing a centralized repository.*

1 Introducción

En los últimos años, la industria de la construcción está creando nuevas soluciones para satisfacer las nuevas necesidades de sus cada vez más exigentes clientes. Entre estos, son los empresarios los que muestran un mayor grado de exigencia, requiriendo soluciones flexibles y eficientes. En consecuencia, la tendencia actual pasa por crear entornos inteligentes en los inmuebles de oficinas, en los cuales los usuarios pueden configurar la iluminación, la temperatura, los permisos de acceso a diferentes infraestructuras e incluso la distribución física del espacio. Son los llamados espacios reconfigurables cuya distribución puede alterarse mediante un control remoto que se comunica con una red control mediante infrarrojos.

Este tipo de solución tiene algunas limitaciones en cuanto a la gestión de preferencias, seguridad, consumo de energía y dimensiones. En primer, lugar el control remoto no permite almacenar las preferencias del usuario y son necesarios reajustes cada vez que se usa. Otro problema tiene que ver con el consumo de energía de los controles remotos cuya batería tiene un tiempo de vida limitado, por lo que es necesario reemplazarla cada cierto tiempo. Además, no debe obviarse que los controles remotos están asociados a espacios, no a usuarios.

En este artículo se presenta un modelo que integra tarjetas inteligentes que almacenan perfiles de usuarios en una red de control para configurar, almacenar y reconfigurar los llamados espacios inteligentes. Con el empleo de tarjeta inteligentes se superan varias de las limitaciones asociadas al empleo de controles remotos ya que son muy pequeñas y no necesitan baterías. Además, son capaces de almacenar preferencias y certificados de usuarios, siendo posible asimismo modificar la información de configuración almacenada de manera sencilla.

Conviene resaltar que los aspectos relacionados con el control de acceso o la autenticación de usuarios no han sido considerados en esta primera etapa. Es por ello por lo que se asume que el usuario que intenta configurar la red ha sido previamente autorizado por un sistema de control anterior. La integración de un módulo de seguridad se encuentra entre los planteamientos de trabajos futuros.

Muchos escenarios diferentes son posibles: un usuario puede controlar con su tarjeta la intensidad de las luces en su casa, la temperatura de las habitaciones, la

Este trabajo ha sido parcialmente financiado por el Ministerio de Industria de España dentro de los proyectos Trust-eS (FIT-070000-2003-930) y Everyware (TIC2003-08995-C02-01), y por la Comisión Europea (VI Programa Marco) dentro del proyecto Ubisec (IST-2002-506926).

posición de las persianas, sus preferencias en cuanto a programas de televisión,....

La solución propuesta se basa en una tecnología de redes de control específica, LonWorks, que proporciona una arquitectura abierta y distribuida para controlar y gestionar cualquier tipo de sensres o actuadores.

La estructura del artículo se presenta a continuación: En la sección 2 se describen las tecnologías Tarjeta inteligentes y LonWorks y además se incluyen una revisión por trabajos relacionados. La sección 3 describe la arquitectura propuesta mientras que la sección 4 se dedica a la definición de perfiles de usuarios. Por último se presentan las conclusiones obtenidas

2 Estado del Arte

En esta sección se revisan las principales características de las tarjetas inteligentes y se describe la tecnología de automatización elegida, LonWorks. Finalmente, se presenta un resumen con varios trabajos y proyectos relacionados.

2.1 Tarjetas inteligentes

Una tarjeta inteligente es un dispositivo capaz de almacenar y procesar información manteniéndola protegida de accesos no autorizados o fraudulentos. Para funcionar, la mayoría de ellas necesitan un lector o CAD (*Card Acceptance Device*) que le proporciona tanto alimentación como una interfaz de comunicaciones.

El mecanismo de comunicación entre una tarjeta inteligente y un lector es definido en el estándar ISO 7816-4 [1] y se basa en el uso de APDUs (*Application Protocol Data Units*) que son paquetes de información con un formato específico. Una tarjeta inteligente nunca inicia la comunicación con el CAD pero es capaz de contestar a comandos enviados por el lector. Se han definido dos tipos de APDU; COMMAND APDU (enviadas a la tarjeta) y las RESPONSE APDU (enviadas por la tarjeta como una respuestas a una COMMAND APDU) (ver Tablas 1 y 2).

Código	Nombre	Long.	Descripción
CLA	Class	1	Clase de instrucción
INS	Instrucción	1	Código de instrucción
P1	Parámetro 1	1	Parámetro 1
P2	Parámetro 2	1	Parámetro 2
Lc	Long.	variable	Nº bytes de datos
Data	Data	Lc	Datos enviados
Le	Long.	variable	Max. Nº de bytes en la respuesta

Tabla 1: Contenido de la Command APDU

En concreto las tarjetas elegidas para el desarrollo han sido las Java Cards Cyberflex 32K de Axalto (antes Schlumberger). Estas tarjetas implementan una versión limitada de Java (JavaCard 2.1.1 [2]) con sus correspondientes API, JCRE y JVM). La principal razón de esta elección radica principalmente en las ventajas que supone el desarrollo en Java y, por ende, en el alto grado de difusión que las tarjetas Java Card tienen entre la comunidad de desarrolladores.

Código	Nombre	Long.	Descripción
Data field	Data	Lr	Campo de datos de la respuesta
SW1	Status byte 1	1	Estado del procesado de comandos
SW2	Status byte 2	1	Cualificador del procesado de comandos

Tabla 2: Contenidos de las APDU de respuesta

2.2 Redes LonWorks

La red LonWorks [3], a diferencia de otras redes de control, es una red descentralizada en la que no existe un elemento central de control. Para asegurar la interoperabilidad entre dispositivos de diferentes fabricantes se creó la asociación LonMark ([4], [5]) con el objetivo de desarrollar perfiles estándar para asegurar ésta mediante la certificación de productos LonWorks y la promoción de los beneficios de los sistemas interoperables.

La arquitectura LonWorks consiste en varios dispositivos inteligentes o nodos que se comunican entre sí mediante un protocolo común, LonTalk [6], sobre diferentes canales de comunicación. Cada nodo en se asocia con una funcionalidad específica en la red y los resultados son compartidos entre todos los nodos instalados.

El protocolo LonTalk define dos maneras de intercambiar mensajes entre dispositivos diferentes. La primera de ellas se basa en el envío y recepción de variables de red (de entrada o de salida) para la transmisión de información entre dispositivos. La segunda manera, que no es implementada por la mayoría de fabricantes, consiste en el uso de mensaje explícitos entre nodos.

Toda red LonWorks requiere un proceso de instalación en el cual se asocia una funcionalidad específica con cada dispositivo en la red programándolo con una aplicación que puede dividirse en uno o varios bloques funcionales. Cada uno de estos bloques implementa un perfil funcional, un conjunto de variables de red y

propiedades de configuración que determinan un comportamiento específico de un dispositivo.

Parámetro	Funcionalidad
Neuron ID	Identificador único de 48-bits para dispositivos LONWORKS
Standard program ID (SPID)	Identificador de la interfaz
Device channel ID	Canal al que está conectado el dispositivo (opcional)
Device location field	Localización del dispositivo (opcional)
Device self-documentation string	Descripción de los bloques funcionales en un dispositivo
Device configuration properties	Información de configuración para el dispositivo y los bloques funcionales
Functional blocks	Componentes lógicos implementados en el dispositivo

Tabla 3 : Elementos de la interfaz externa

El paso final de este proceso implica la creación de una interfaz externa para cada dispositivo, a través del cual el resto de nodos puedan conocer su funcionalidad y viceversa. Dicha interfaz está formada por los elementos que se muestran en la tabla 3.

Una vez finalizado, este proceso es posible la asignación de variables de red a los nodos conocida como *binding* o enlazado. Gracias a este proceso se establece un cable virtual entre los dispositivos de manera que cuando una variable de red modifica su estado o valor, el cambio es transmitido a todos los dispositivos con los que esté enlazado.

Es posible añadir diferentes atributos a las variables de red para proveer utilidades adicionales. De esta forma las variables de red pueden establecerse como autenticada (se emplean mensajes autenticados para transmitir sus valores), con prioridad (cuando se emplean time slots de prioridad para transmitir sus valores) y síncronos (todos los valores asignados a una variable de red son propagados). Las variables de red junto con algunas propiedades de configuración se agrupan generalmente en perfiles funcionales para llevar a cabo una función en un dispositivo.

2.3 Trabajos relacionados

Hasta ahora, las tarjetas inteligentes solo se habían combinado con redes de automoción, bien en un entorno industrial o de oficina como un simple medio de autenticación de usuarios, como por ejemplo ocurre en los proyectos FUTURE HOME [7] o ePerSpace [8]. La aproximación más cercana a la propuesta en este artículo se puede encontrar en un *whitepaper* de SUN [9] donde una visión más ambiciosa de la personalización de entornos con perfiles de usuarios se presenta como posible escenario. Así, en dicho documento se describe como podría ser posible emplear perfiles de usuarios en un "hogar conectado" para configurar la red en función de las preferencias del usuario. Lamentablemente no se trata de una propuesta técnica si no de un ejemplo para justificar la ida de redes residenciales basadas en una pasarela.

3 Propuesta de interoperabilidad

En este artículo se propone una arquitectura, ilustrada en la Figura 1, para configurar una red de control, en particular una red LonWorks, con una tarjeta inteligente. Por lo tanto, el elemento clave de este diseño será la pasarela, un nodo LonWorks dedicado que actúa como un puente entre la red LonWorks y un lector de tarjeta inteligentes. Es responsable de interpretar la información almacenada en la tarjeta para configurar la red y debe ser capaz de programar una tarjeta con la configuración asociada al entorno donde está instalado.

Una red de control está diseñada para llevar a cabo una funcionalidad concreta e invariable durante un largo periodo de tiempo. Dentro del tiempo de vida de una red de control, se puede distinguir una primera fase de instalación donde la red se instala y configura, y una segunda de operación, en la cual la red ya está en funcionamiento y donde se comporta según las funcionalidades definidas en la fase anterior.

Debido a este modo de funcionamiento, la pasarela no va a poder procesar cualquier tipo de información almacenada en la tarjeta, sino que únicamente va a ser capaz de responder ante la información que tiene definida en su interfaz de red externa.

Por lo tanto, la pasarela debe proporcionar dos modos de funcionamiento distinto: por un lado, programar una nueva tarjeta con los perfiles almacenados en la red (Modo Configuración) y, por otro lado, recibir la información de una tarjeta que ha sido programado previamente y configurar la red con la información que se recibe de la tarjeta (Modo Operación).

LonMark ha estandarizado distintos perfiles para realizar funciones de control muy diversas. A modo de ejemplo, basta con citar la variedad de perfiles definidos para controlar las luminarias de un edificio o vivienda (*Light Sensor, Switch, Manual Override Switch, Constant Light Controller, Lamp …*).

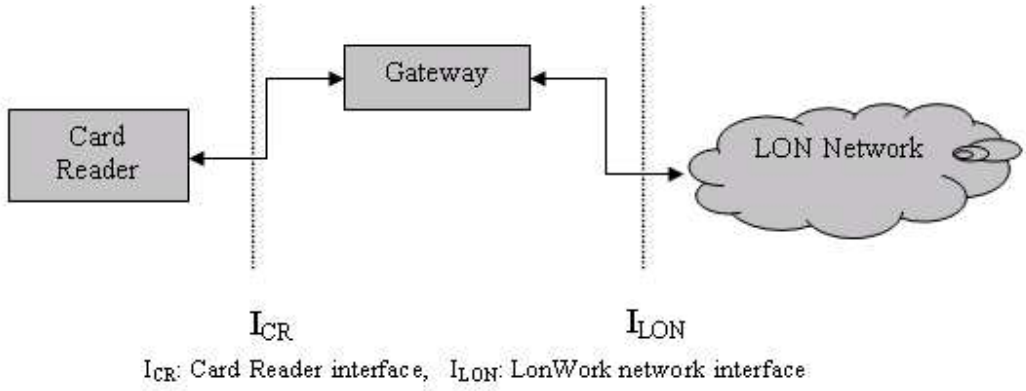


Figura 1: Arquitectura para integrar la tarjeta inteligente en una red LonWorks

En cada uno de estos perfiles se especifica una funcionalidad muy concreta que se implementa definiendo, dentro de cada dispositivo, un conjunto de variables (*Standard Network Variables Types* o *SNVT* y *Standard Configuration Parámetro Type* o *SCPT*) y en asociar un determinado comportamiento a cada una de ellas. Las variables de tipo *SNVT* se utilizan para transmitir información entre distintos dispositivos LonWorks, mientras que las variables de tipo *SCPT* se utilizan para configurar parámetros asociados a cada uno de los perfiles.

De este modo, una red que controle la iluminación de una vivienda y que tenga implementado alguno de estos perfiles, va a ser capaz de reaccionar ante distintos eventos (transición día-noche, incremento de la luminosidad al abrir una persiana...). Dependiendo del perfil o de los perfiles implementados, la reacción a estos eventos podrá ser ajustada tanto de forma manual como automática en función del estado de las variables de tipo SCPT.

De esta forma la personalización de la iluminación de una habitación, una sala,... a través de una tarjeta inteligente podría realizarse almacenando las condiciones antes las cuales la red va a reaccionar. Concretando un poco más, esto implicaría actualizar, al menos, dos variables por luminaria: la variable de tipo *SNVT* red que selecciona el nivel de intensidad de la luminaria (*nviSetting*) y la variable de tipo *SCPT* que modifica el umbral de encendido de la luminaria. (*nciLuxSetPoint*).

Es importante resaltar que el tipo de cada variable en la pasarela y el de sus asociadas en la red deben coincidir y que, durante el proceso de instalación de la red LonWorks, deben ligarse con el nodo o nodos que implementen el *Constant Light Profile*. Ello implica que, en el modo Configuración, la pasarela va a enviar a la tarjeta tanto las variables de red como sus valores asociados en ese instante.

Esto lleva a otra consideración de diseño importante, el formato de la información que es intercambiada entre la red y la tarjeta y, posteriormente, almacenada en esta última. A lo largo de las secciones siguientes diferentes estructuras lógicas empleadas para almacenar la información de red son mostradas.

3.1 Propuesta de interoperabilidad

La solución propuesta se puede considerar como "semiindependiente" de la red ya que en la tarjeta inteligente se almacenan estructuras de datos que se refieren a una red concreta y que reúne un conjunto de registros con los cuales es posible establecer la configuración de la red. Cada estructura de datos es referenciada por un identificador, el identificador de entorno, que debe representar de manera única la red.

Cada uno de los registros anteriormente descritos contiene el nombre de un elemento de red a configurar, su valor e información a cerca del periodo de validez del mismo. En esta primera aproximación, la variable de red a controlar y su valor en Neuron C [10] (el lenguaje de programación para dispositivos LonWork) son almacenados en los campos nombre y valor. De esta forma, existe un grado de dependencia intermedio entre la información almacenada en la tarjeta y la red de control ya que se ha definido una estructura general para almacenar la información de red pero se ha aplicado a una red de control concreta.

3.2 Identificador de entorno

El identificador de entorno, como su nombre indica, permite identificar una red LonWorks concreta y es definido durante el proceso de instalación de la red. Cuando una tarjeta es insertada en el lector la pasarela le envía dicho identificador.

3.3 Comunicación entre la tarjeta inteligente y la pasarela

Las funcionalidades de gestión asignadas a la tarjeta inteligente son llevadas a cabo por una aplicación almacenada en la misma. Dicha aplicación, denominada, CMA (*Configuration Manager Application*), es capaz de comunicarse con la red LonWorks y de gestionar las diferentes configuraciones de usuario.

Una vez que la tarjeta es insertada, la pasarela debe seleccionar la aplicación CMA correspondiente entre las que pueda haber almacenadas en la tarjeta. Entonces será posible establecer una comunicación entre tal aplicación y la que se ejecuta en la pasarela. La selección de la aplicación CMA se realiza mediante el envío de la APDU mostrada en la Tabla 4. Dicha APDU es estándar y se usa para seleccionar en la tarjeta una determinada aplicación identificada por su AID que no es más que una secuencia de entre 5 y 16 bytes que identifica de manera única la aplicación.

CLA	INS	P1	P2	Long.	Body
00	A4	04	00	AID Long.	AID

Tabla 4: Formato de la APDU SELECT empleada para seleccionar una aplicación

La respuesta normal a esta APDU es otra de respuesta con el valor 0x9000 cuando el proceso de selección de aplicación culmina con éxito (ver Tabla 5); en caso contrario una APDU de error será envíada a la tarjeta inteligente.

SW1	SW2	Data field
90	00	-

Tabla 5: APDU de respuesta para operación exitosa.

Tras seleccionar el CMA, la pasarela le enviará el identificador de entorno mediante el comando APDU representado en la tabla 6.

CLA	INS	P1	P2	Long.	Body	Le
AA	01	0?	00	10	Env. ID	--

Tabla 6: APDU empleada para enviar el identificador de entorno.

De nuevo, la respuesta a este comando será una APDU con valor 0x9000 en caso de éxito y una respuesta de error, en la que se indica la razón del mismo empleando un código numérico. Dependiendo en el modo en el que la pasarela está trabajando (como se muestra en 4) esta APDU solicitará al CMA la creación de una configuración de usuario en la que almacenar las variables de red o indicará a la tarjeta que las variables de red asociadas con el identificador de entorno son requeridas.

4 Modos de operación de la pasarela

En esta sección se describe en detalle la interacción entre una tarjeta inteligente (una CMA almacenada en ella en realidad) y la pasarela tanto en el modo de Configuración como en el de Operación. El modo de Configuración es necesario para proporcionar a la tarjeta un conjunto de perfiles personalizados para controlar y configurar la red. En cambio, en el modo de Operación la pasarela opera como un traductor entre la tarjeta inteligente y la red.

4.1 Modo Configuración

En este modo, la pasarela envía a la tarjeta inteligente la configuración actual de la red representada por el identificador de entorno y su interfaz externa (las variables en la red que son controladas por él). En principio se considera que se almacenan en la tarjeta inteligente todos los pares variable de red – valor existentes en la red. En particular deben enviarse el conjunto de variables de red de la pasarela que se encuentran ligadas a variables en otros dispositivos. Las arrays de variables de red, se enviarán elemento a elemento como si estos fuesen una variable de red aislada, incluyendo el número del elemento en el array como parte del nombre.

En este modo, tras seleccionar la aplicación CMA y enviarle el identificador de red, dicha aplicación puede crear una estructura de datos para almacenar la configuración de red. Entonces, la pasarela envía todas las variables de red (*Network Variables, NVs*) definidas en su interfaz externa a la tarjeta. Continuando con el ejemplo de control de iluminación, la pasarela tiene que enviar a la tarjeta inteligente dos variables por cada dispositivo o conjunto de dispositivos bajo su control. En ejemplo, la variable de red *nviSetting* y la propiedad de configuración *nciLuxSetPoint* con sus valores correspondientes. Es importante resaltar que previamente al envío la red y los dispositivos deben configurarse en el estado que desea el usuario ya que dichos valores serán los almacenados en la tarjeta inteligente.

Todas las transacciones son confirmadas de manera que el CMA responde con un OK si es capaz de almacenar la información recibida. Al final de la secuencia, la tarjeta inteligente habrá recibido y almacenado toda la información necesaria para configurar y controlar la red (ver Figura 2).

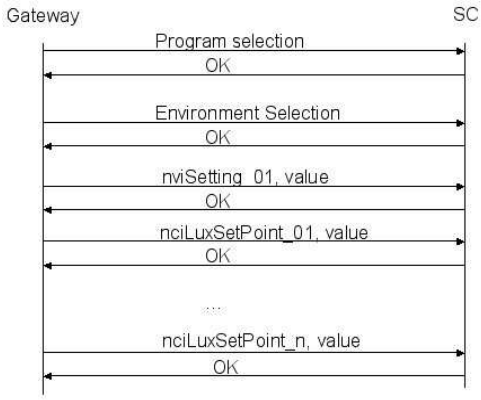


Figura 2: Secuencia de mensajes del modo configuración

Si se produce un error durante este proceso una respuesta de error será enviada a la pasarela quien interrumpirá la comunicación. En general se identifican tres tipos diferentes de errores:

- *Error de Aplicación*. Debido a una selección errónea de la aplicación en la tarjeta bien porque ésta no existe bien porque no está preparada para hablar con la pasarela.
- *Error de Entorno.* Este error se produce cuando no existe una configuración para el identificador de entorno recibido. Cuando ocurre este error la pasarela puede pasar automáticamente a crear dicha configuración en la tarjeta.
- *Error de Petición*. Este error se produce cuando la aplicación no es capaz de crear la estructura de datos en la tarjeta porque no hay memoria suficiente para ello.

En cualquier caso, sea cuál sea la causa del error, la pasarela responde a este cargando una configuración segura definida por defecto a modo de "*fail-safe*". De este modo se evita que la red quede parcialmente configurado o en un estado peligroso o inestable.

4.2 Modo de operación

Cuando la pasarela funciona en este modo está preparada para recibir instrucciones de la tarjeta para configurar la red LonWorks. El proceso de configuración comienza cuando la tarjeta es insertada en lector y se ilustra en la Figura 3. Cuando el CMA recibe una APDU pidiendo la estructura de datos asociada con un identificador de red responde enviando una APDU por cada variable, incluyendo su valor almacenado. Estas APDUs tendrán el valor 0x91 en el campo SW1 si existe una variable siguiente o 0x90 en caso contrario. De esta forma la aplicación en la pasarela puede detectar si la comunicación de las variables ha finalizado. En caso de que no exista variables asociadas con dicho entorno se envía a la pasarela una APDU de error.

Una vez que la pasarela recibe una variable de red desde la tarjeta, debe comprobar que el elemento de red referenciado en la APDU ha sido previamente declarado en su interfaz externa. De otro modo, debe descartarse esa APDU y finalizar la comunicación.

Para configurar la iluminación, en el ejemplo inicial, el usuario debe enviar el conjunto de variables de red almacenadas para esa red entre las que se encuentran las asociadas a tales dispositivos. Esto implica enviar las variables y los valores almacenados anteriormente en el modo Configuración.

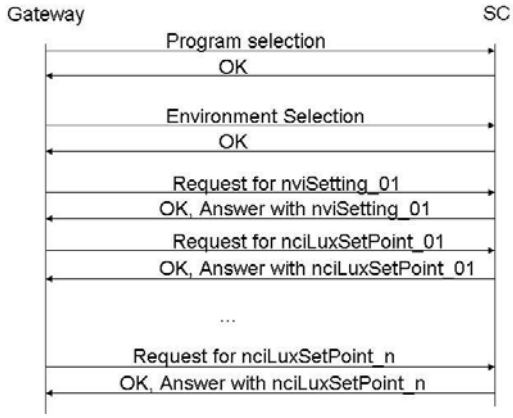


Figura 3: Secuencia de mensajes del modo operación

5 Conclusiones y trabajo futuro

La arquitectura propuesta supone una primera aproximación a un entorno inteligente capaz de adaptarse a las preferencias del usuario. Dichas preferencias se almacenarán en su tarjeta personal y permitirán configurar la red una vez se inserta en el lector.

En la siguiente fase de desarrollo se pretende la definición de nuevas estructuras de configuración que permitan extender el grado de funcionalidad. Además, se llevará a cabo un estudio del rendimiento alcanzado al incrementar el número de operaciones de gestión asignadas a la tarjeta inteligente. También se realizará un análisis para evaluar las ventajas e inconvenientes de distribuir el coste computacional entre la tarjeta. Por un último, se analizará la integración de esta arquitectura con pasarelas residenciales como, por ejemplo, pasarelas OSGi [11].

Referencias

[1] International Organisation for Standardisation (ISO), JTC 1/SC 17. ISO/IEC 7816. Identification cards - Integrated circuit(s) cards with contacts, 1995

[2] Sun Microsystems, Inc. Java Card™ Specifications Version 2.2.1 (2003)

[3] Echelon Corporation. Introduction to the LonWorks System, 1999. Disponible en www.echelon.com/support/documentenion/manuals/078-0183-01A.pdf

[4] LonMark Association. LONMARK Application Layer Interoperability Guidelines version 3.3, 2002. Disponible en www.lonmark.org/products/guides.htm

[5] LonMark Association. LONMARK Layer 1-6 Interoperability Guidelines version 3.3, 2002. Disponible en www.lonmark.org/press/download/LmPhy33.pdf

[6] Echelon. ANSI/EIA709 Control Network Protocol Specification (LonTalk). (CEA-709.1).2002. Disponible en www.echelon.com/products/lonworks/protocol

[7] Juha Koivisto. Wireless Home Network Architecture and Concepts for User Interactions. Technical Report of the Project FUTURE HOME. (IST 2000-28133), 2000

[8] WP3 contributors, G. Nygreen, P. Plaza. Detailed workplan and specification for intermediate test beds. Technical Report of the Project ePerSpace – IST Integrated Project (N° 506775), 2004

[9] SUN Microsystems. The Connected Home Powered by Java EmbeddedServer™. 2001. Disponible en java.sun.com/products/consumerembedded/whitepapers/dot.com.home. pdf

[10] Echelon Corporation. Neuron C Reference Guide Revision 5, 2003. Disponible en www.echelon.com/support/documentation/manuals/078-0140-02E.pdf

[11] The Open Services Gateway Initiative. OSGi Service Platform Release 3 March 2003OSGi Service Platform Release 3. 2003. Disponible en OSGi Official Web Site. www.osgi.org

Segmentación temporal de tráfico fractal mediante transformadas wavelet redundantes y teoría de la información

David Rincón*, Flaminio Minerva†, Sebastià Sallent*, Michele Pagano†
*Departamento de Ingeniería Telemática (enTel), Universidad Politécnica de Catalunya
† Dipartimento di Ingegneria dell' Informazione, Università di Pisa
Escuela Politécnica de Castelldefels (EPSC). Av. Canal Olímpic, s/n.
08860 Castelldefels (Barcelona)
Teléfono: 93 413 70 56. Fax: 93 413 70 07
E-mail: drincon@mat.upc.edu

***Abstract**. Network traffic exhibits fractal characteristics, such as self-similarity and long-range dependence. Traffic fractality and its associated burstiness have important consequences for the performance of computer networks. There are several estimators of the fractal parameters, and those based on the discrete wavelet transform (DWT) are the best in terms of efficiency and accuracy. The DWT estimator does not consider the possibility of changes to the fractal parameters over time. We propose using the Schwarz information criterion (SIC) to detect changes in the variance structure of the wavelet decomposition and then segmenting the trace into pieces with homogeneous characteristics for the Hurst parameter. The procedure can be extended to the stationary wavelet transform (SWT), a non-orthogonal transform that provides higher accuracy in the estimation of the change points. The DWT-SIC and SWT-SIC algorithms have been tested against synthetic and well-known real traffic traces, with promising results.*

1 Introducción

Se ha documentado extensamente que el tráfico transportado por las redes de ordenadores presenta unas características estadísticas que no siempre coinciden con los modelos markovianos y poissonianos (o de dependencia a corto plazo) que tradicionalmente se han utilizado en los estudios de caracterización y dimensionado de las redes de conmutación de paquetes. Este hecho ha llevado al desarrollo de nuevos modelos estocásticos que tienen en cuenta las características fractales del tráfico real, entre las que destacan la autosimilitud (*self-similarity)* y la dependencia a largo plazo (LRD, *long-range dependence*) [1,2].

Se han realizado muy pocos estudios relacionados con el seguimiento de la evolución temporal de los parámetros fractales [3,4]. Es natural esperar que las características fractales del tráfico transportado por redes reales sean variables con el tiempo. Detectar los instantes de cambio e identificar las regiones temporales en las que el carácter fractal es homogéneo puede ser útil para algunos algoritmos que explotan las propiedades de memoria a largo plazo del tráfico; nos estamos refiriendo, por ejemplo, a la modificación del control de congestión de TCP descrita en [5], el control predictivo de tráfico para fuentes de vídeo presentado en [6], o el estimador de ancho de banda efectivo descrito en [7]. Además de estas posibles aplicaciones nuestro estudio se justifica por la necesidad de conocer mejor las características del tráfico con el objetivo de generar trazas sintéticas para probar equipos o realizar simulaciones. Nuestro objetivo es, por tanto, seguir de manera precisa la evolución temporal de los parámetros fractales, lo que nos permitirá conocer mejor la dinámica de la red.

Entre los algoritmos de análisis de la fractalidad los mejores en términos de eficiencia computacional, precisión y flexibilidad son los basados en la Transformada Wavelet Discreta (DWT, *Discrete Wavelet Transform*). El estimador desarrollado por Abry y Veitch [8] es reconocido como el más eficiente y completo de los estimadores de LRD, ya que es capaz de realizar una estimación conjunta de los dos parámetros fractales (α y c_f). La DWT descompone la serie temporal bajo estudio en un análisis multiresolución a diferentes escalas, que convierte el espectro del tipo 1/f, típico de los procesos fractales, en una cierta distribución de la varianza a lo largo de las diversas escalas de análisis. Si el parámetro de escalado cambia en un determinado instante, este cambio se produce simultáneamente en todas las escalas. Por tanto, una manera bastante lógica de detectar el cambio en el parámetro fractal es monitorizar cada una de las escalas temporales (que no son más que las salidas de la transformada wavelet) y aplicar un algoritmo de detección de cambios de varianza a cada escala. La detección de un cambio simultáneo en todas las escalas (o en un número significativo de ellas) señalará un cambio global en las características fractales del tráfico. Mediante una fase final de *clustering* y decisión se determina si el candidato a

Este trabajo ha sido apoyado por el Ministerio de Educación y Ciencia (TIC2003-09042-C03-02) y FEDER, y se realiza en el marco de la red europea de excelencia EuroNGI, financiada por la Comisión Europea (project nº 507613).

punto de cambio es real o si debe ser descartado y continuar la monitorización.

En este estudio proponemos el uso del Criterio de Información de Schwarz (SIC, *Schwarz Information Criterion*) como algoritmo de detección de cambios, aplicado a la salida de la transformada wavelet. SIC se basa en el cálculo de la función de máxima verosimilitud de las hipótesis nula (no hay cambio en la secuencia) y alternativa (existe algún cambio).

Los aspectos más deficientes de los estimadores basados en DWT son la falta de precisión a escalas temporales elevadas (correspondientes a frecuencias bajas, que es donde la LRD se hace patente), y su incapacidad para seguir la evolución temporal de los parámetros fractales. La DWT no es la única transformada capaz de capturar el carácter fractal del tráfico. Una alternativa es la Transformada Wavelet Estacionaria (SWT, *Stationary Wavelet Transform*), una transformada no ortogonal que genera muestras temporales adicionales, lo que nos permite localizar con más precisión los puntos de cambio.

Este artículo presenta los algoritmos DWT-SIC y SWT-SIC. En la siguiente sección se presenta brevemente la LRD y su estimación con la DWT; en la tercera sección se presenta el SIC. La cuarta sección describe cómo funcionan conjuntamente las transformadas y el criterio, junto con las validaciones realizadas con trazas sintéticas, los resultados con trazas reales y algunos apuntes sobre una versión en tiempo real de los algoritmos. El artículo finaliza con las conclusiones y las líneas actuales de trabajo.

2 LRD y su estimación mediante la transformada wavelet

2.1 Descripción de la LRD

Un proceso estacionario x(t) presenta dependencia a largo plazo (LRD) si su autocorrelación decae a un ritmo menor que el de una exponencial negativa. De manera equivalente, la LRD se puede definir también a partir del comportamiento del espectro en el origen:

$$S_x(f) \sim \frac{c_f}{|f|^{\alpha}} \text{ para } |f| \to 0 \qquad (1)$$

Los parámetros de la LRD son α (denominado parámetro de escalado, y que viene a ser una medida de la intensidad de la fractalidad) y c_f (que tiene unidades de varianza y está relacionado con el volumen del tráfico). El parámetro de escalado se suele expresar mediante $H = (1+\alpha)/2$, el denominado "parámetro de Hurst". Tradicionalmente se ha prestado más atención a H que a c_f, pero la importancia de este último no es despreciable, ya que tiene una influencia notable en las pérdidas producidas en una cola cuando se inyecta tráfico LRD, y también aparece en la expresión de la varianza de la estimación de la media de un proceso LRD, determinando por tanto los intervalos de confianza en la estimación. En el resto del artículo centraremos nuestra atención en α y H, aunque los algoritmos descritos también pueden monitorizar c_f.

2.2 La transformada wavelet discreta

La DWT es una potente herramienta utilizada desde hace tiempo en el campo del procesado de señal, que permite una estimación precisa, rápida y eficiente de la LRD presente en el tráfico. Dada una señal x(t), los coeficientes $d_x(j,k)$ de la transformada wavelet de x(t) a la escala j y en tiempo k se definen como el producto escalar $d_x(j,k) = \langle x, \psi_{j,k} \rangle$, donde la base de la descomposición es una familia de funciones construida a partir de la traslación y escalado de una función denominada "wavelet madre" ψ_0:

$$\psi_{j,k}(t) = 2^{-j/2}\psi_0(2^{-j}t - k), \text{ j} = 1\ldots \text{J}, \text{k} \in \text{Z}. \qquad (2)$$

La DWT puede entenderse como el resultado del filtrado de la señal a través de una banco de filtros espejo en cuadratura que descomponen la señal en una subbanda paso alto y otra paso bajo, seguidos por un diezmado necesario para asegurar la ortogonalidad de la descomposición. El proceso se itera sobre la señal paso-bajo hasta el nivel J deseado. A la salida se obtiene una descomposición en subbandas diádicas, en la que la primera componente ocupa la mitad superior del espectro, la segunda ocupa la cuarta parte del espectro inmediatamente inferior, y así sucesivamente. Este proceso da lugar a un análisis multiresolución en el que la señal inicial es descompuesta en una aproximación paso-bajo a la escala J, $a_x(J,k)$ y un conjunto de detalles paso-alto $d_x(j,k)$ para cada escala j = 1..J.

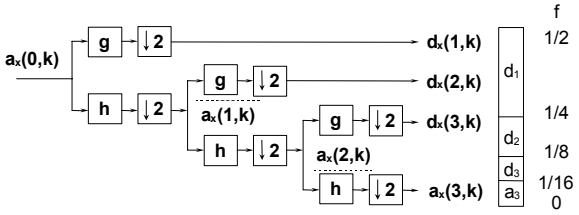


Figura 1: banco de filtros DWT en una descomposición de nivel 3. Derecha: descomposición en subbandas del espectro normalizado.

2.3 Análisis de procesos LRD con la DWT

Abry y Veitch desarrollaron el diagrama LogScale [8], un estimador sin sesgo y computacionalmente eficiente de los parámetros de la dependencia a largo plazo. Se basa en el cálculo de la varianza μ_j de los coeficientes wavelet de cada subbanda j, que se puede interpretar como la potencia de la señal en dicha subbanda. A partir de la expresión (2),

$$\mu_j = E\left[d_x^2(j,k)\right] = 2^{j\alpha} \cdot c_f C(\alpha, \psi_0) \qquad (3)$$

y tomando logaritmos a ambos lados,

$$\log_2(\mu_j) = j\alpha + \log_2(c_f C) + g_j \qquad (4)$$

donde g_j es un factor de corrección de sesgo (necesario porque el logaritmo de la esperanza no coincide con la esperanza del logaritmo) que sólo depende de n_j (número de coeficientes wavelet a escala j). Los parámetros α y $c_f C$ (nótese la dependencia de C con α) pueden ser estimados a partir de (3) mediante una regresión lineal ponderada, donde el peso de cada muestra de la regresión es proporcional a n_j (a más coeficientes, más fiable es la estimación de la varianza y por tanto más peso tendrá esa escala en la regresión lineal). Asumiendo una distribución normal de los coeficientes se obtienen los intervalos de confianza de las estimaciones.

En la figura 2 se puede observar un ejemplo de uso del diagrama LogScale, aplicado a una de las trazas más utilizadas en la literatura: pAug89, correspondiente a la captura de un millón de tramas en la red Ethernet de Bellcore, en agosto de 1989, y estudiada detalladamente en [1]. En este caso hemos agregado las bytes llegados en slots de 10 ms, reduciendo la traza a 314282 muestras[2]. Se observa alineación de los logaritmos de las varianzas. La regresión lineal entre las escalas j_1=5 y j_2=17 devuelve α = 0.605 (0.585, 0.626) y H = 0.803, siendo (0.792, 0.813) el intervalo de confianza del 95%.

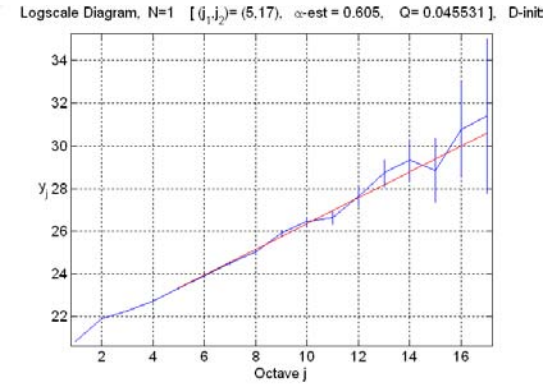


Figura 2: Ejemplo de análisis de la traza pAug89.

Las mayores desventajas del diagrama LogScale son la falta de precisión a altas escalas (las subbandas más bajas, donde la transformada obtiene pocos coeficientes, pero que son la zona del espectro donde la LRD se hace patente) y su falta de adaptación a cambios temporales (el análisis se hace sobre una traza completa). Existe una versión progresiva del estimador [3] pero se trata de un análisis acumulativo, no dinámico (es decir, se puede calcular el "valor medio" de los parámetros LRD a partir de las muestras acumuladas desde t=0). Los mismos autores desarrollaron un test estadístico para determinar la estacionariedad de los parámetros LRD [4], pero sólo puede detectar cambios en segmentos diádicos (de longitud 2^n). Esto es claramente insuficiente, por lo que nos hemos puesto como objetivo el estudio de transformadas wavelet alternativas para solucionar las carencias de la DWT.

[2] La elección de la escala de 10 ms se justifica por ser la más baja en que podemos eliminar casi totalmente los slots con 0 bytes.

2.4 Transformada wavelet estacionaria

La SWT es muy similar a la DWT, diferenciándose en la desaparición de la fase de diezmado a la salida de los filtros. Por tanto la SWT es una transformada redundante que dobla, aproximadamente, el número de coeficientes a la salida a cada iteración del filtrado. Este hecho proporciona una precisión más elevada en el cálculo de las varianzas de cada escala (debido a la mayor cantidad de muestras) y lo que es más importante, abre la puerta a una caracterización de la evolución de la señal en el dominio temporal. Por otro lado, el cálculo de la SWT requiere mucha más memoria, y el coste de procesado posterior de los coeficientes también se incrementa notablemente.

La expresión para la estimación de los parámetros LRD es similar a (4), con una diferencia en el término constante debido a la redundancia. Los términos de corrección de sesgo son esencialmente los mismos que para la DWT, excepto que en este caso su valor es constante para todas las escalas (recordemos que sólo dependían de n_j, el número de coeficientes, que es en este caso constante para todas las escalas). Los intervalos de confianza gaussianos también son constantes, produciendo una estimación de peso constante a lo largo de todo el diagrama LogScale, lo que simplifica enormemente el coste computacional del proceso. Pero la mayor ventaja es un incremento de la precisión a las escalas altas, donde la DWT proporciona pocos coeficientes.

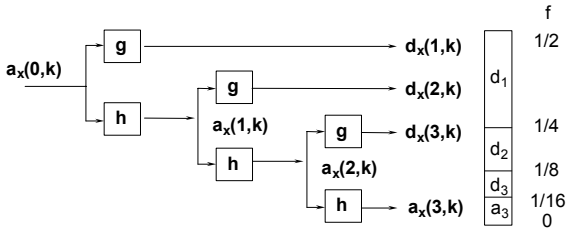


Figura 3: Izquierda: banco de filtros para una SWT de nivel 3. Derecha: descomposición en subbandas, similar a la de la DWT.

La DWT no es capaz de seguir la evolución temporal de la señal con la misma precisión a todas las escalas, debido a su estructura de descomposición del plano tiempo-frecuencia (las escalas inferiores tienen más coeficientes y más resolución temporal pero menos frecuencial, y viceversa para las escalas altas). Puesto que cada subbanda de la SWT tiene la misma cantidad de coeficientes, los instantes de transición de la estructura de varianzas pueden ser detectados con fiabilidad a todas las escalas. Este hecho abre la puerta a un estimador de la evolución temporal de H.

En trabajos anteriores estudiamos el uso de la estadística ICSS (*Iterative Cumulative Sum of Squares*) [9]. Los resultados de la ICSS son buenos, pero le falta flexibilidad a la hora de calcular el valor crítico (calculado mediante el método MonteCarlo) y es difícil de implementar en modo secuencial (en el sentido de progresiva). El Criterio de Información de Schwarz (SIC) es una alternativa sólida.

3 Detección de cambios de varianza

3.1 Planteamiento del problema

Dada una secuencia x_1, x_2 ... x_n de variables aleatorias gaussianas independientes con una media común μ y varianzas σ_1^2, σ_2^2... σ_n^2, probamos la hipótesis nula

$$H_0: \quad \sigma^2_1 = \sigma^2_2 = \ldots = \sigma^2_n = \sigma^2 \qquad (5)$$

contra la hipótesis alternativa

$$H_1: \quad \sigma^2_1 = \ldots = \sigma^2_{k1} \neq \sigma^2_{k1+1} = \ldots = \sigma^2_{kq} \neq \sigma^2_{kq+1} \ldots = \sigma^2_n \qquad (6)$$

donde q es la cantidad (desconocida) de puntos de cambio, y $1 \leq k_1 < k_2 < \ldots < k_q < n$ son sus posiciones. Siguiendo el procedimiento de segmentación binaria sugerido en [10], se puede reducir el problema a la búsqueda de un solo punto de cambio, mediante la iteración del proceso en las dos subsecuencias que rodean la transición. Si se detecta un cambio en una de las subsecuencias, se divide de nuevo y se itera el proceso hasta que no se encuentran más cambios. Por tanto el problema se reduce a probar la alternativa nula contra

$$H_1: \quad \sigma^2_1 = \ldots = \sigma^2_{k0} \neq \sigma^2_{k0+1} \ldots = \sigma^2_n \qquad (7)$$

3.2 Criterio de Información de Schwarz

Una de los algoritmos más utilizados para detección de puntos de cambio es el Criterio de Información de Akaike (AIC) para selección de modelos [11]. A partir del trabajo de Akaike aparecieron variantes que aplican la teoría de la información en otros campos. Un ejemplo es el Criterio de Información de Schwarz [12], definido como sigue:

$$SIC(k) = -2\log L(\hat{\theta}) + p\log n \qquad (8)$$

donde $L(\hat{\theta})$ es el máximo de la función de verosimilitud del modelo, p es el número de parámetros libres, y n es la longitud de la secuencia. En nuestro problema tenemos dos modelos, correspondientes a las dos hipótesis. La decisión sobre la existencia del punto de cambio se toma siguiendo el principio de mínima información: no rechazamos H_0 si $SIC(n) \leq \min_k SIC(k)$, y se rechaza H_0 si $SIC(n) > SIC(k)$ para alguna posición k, y se estima la posición del punto de cambio $\hat{k}$ tal que

$$SIC(\hat{k}) = \min_{1 \leq k < n} SIC(k) \qquad (9)$$

donde SIC(n) es el valor de SIC bajo H_0 y SIC(k) es el valor de SIC bajo H_1 para k=1….n-1. La expresión (8) puede ser reescrita como:

$$SIC(n) = n\log 2\pi + n\log\hat{\sigma}^2 + n + \log n \qquad (10)$$

$$SIC(k) = n\log 2\pi + k\log\hat{\sigma}_1^2 + (n-k)\log\hat{\sigma}_2^2 + n + 2\log n \qquad (11)$$

donde

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2,\ \hat{\sigma}_1^2 = \frac{1}{k}\sum_{i=1}^{k}(x_i-\mu)^2,\quad \hat{\sigma}_2^2 = \frac{1}{n-k}\sum_{i=k+1}^{n}(x_i-\mu)^2$$

son los estimadores sesgados de la varianza de la secuencia entera y las dos subsecuencias 1...k y k+1...n, respectivamente. La capacidad de detección se limita a $2 \leq \hat{k} \leq n-2$, por lo que (9) se modifica:

$$SIC(\hat{k}) = \min_{2 \leq k \leq n-2} SIC(k) \qquad (12)$$

La idea intuitiva tras las matemáticas del criterio SIC es la siguiente: una secuencia sin cambios de varianza es ordenada, tiene poca entropía o información, mientras que una secuencia con uno o más cambios de varianza es más rica en entropía. Por tanto, si dos subsecuencias 1...k y k+1...n de la secuencia original presentan menos entropía que la secuencia entera, es porque la secuencia tiene un cambio de varianza en k.

En [10] se demuestra que el criterio SIC proporciona una estimación consistente del punto de cambio. También se dispone de una expresión para el cálculo del nivel de significación de la estimación, a través del nivel crítico c_α, que modifica nuevamente (12): ahora el criterio será aceptar H_0 si SIC(n) es menor que el mínimo de SIC(k) + c_α para alguna $\hat{k}$, y dicha posición será la estimación del punto de cambio.

La expresión del nivel crítico se deriva de la distribución asintótica de la estadística SIC, y se puede consultar en [10]. En la misma referencia se pueden encontrar otras versiones del estimador, como por ejemplo la versión sin sesgo y el estimador para el caso en que la media es desconocida. El primero proporciona un pequeño incremento en la precisión de la estimación, a costa de una mayor complejidad de cálculo. El segundo no es necesario para nuestra aplicación, ya que la media de los coeficientes wavelet de cada escala es siempre cero. En el presente estudio utilizaremos la definición inicial, modificada con el nivel crítico c_α.

4 Algoritmos DWT-SIC y SWT-SIC

4.1 DWT-SIC

La principal aportación de nuestro trabajo es la conexión entre la salida de la transformada wavelet y los algoritmos de detección de cambio de varianza, que son aplicados a las secuencias que se obtienen en cada una de las ramas de la transformada; es decir, a cada una de las escalas temporales del análisis multiresolución. Si detectamos un cambio simultáneo en todas, o al menos en un conjunto significativo de escalas, dicha posición señalará un punto de cambio en el parámetro de escalado o de Hurst.

A continuación describiremos una prueba realizada con una traza sintética que incluye una transición entre dos segmentos de ruído gaussiano fraccional (FGN, *Fractional Gaussian Noise*) con parámetros de Hurst H=0.5 (LRD nula) y H=0.9 (LRD alta). Los dos fragmentos se han sintetizado mediante el método presentado en [17], y son independientes. En la figura 4 se puede observar el cambio de la varianza a todas las escalas. La única excepción es la escala 3, sobre la cual pivota la basculación de las dos rectas. Dichos fenómenos, que llamaremos “puntos ciegos”

aparecen en todas las transiciones, y corresponden a la escala (o escalas, ya que a veces la ceguera se extiende a dos escalas) en la que la varianza a uno y otro lado de la transición es prácticamente idéntica; es decir, a la escala en la que las regresiones lineales de la alineación de varianzas se cruzan. El adjetivo "ciego" nos indica que los algoritmos de cambio de varianza nunca detectarán el cambio a dicha escala.

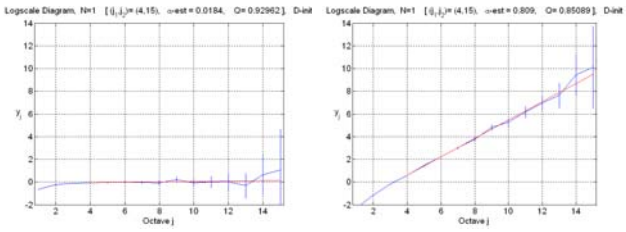


Figura 4: Diagramas LogScale para una secuencia FGN con H=0.5 (izquierda) y con H=0.9 (derecha).

El diezmado realizado en cada proceso de filtrado de la DWT implica que el número de muestras disponibles en cada escala sea diferente. Por ejemplo, si se utiliza la wavelet de Haar y la longitud de la secuencia de entrada es de 1024 muestras, la primera rama de la wavelet (la correspondiente a la subbanda de frecuencia más elevada) obtendrá 512 coeficientes d_1, la segunda rama obtendrá 256 coeficientes d_2, y así sucesivamente (para otras wavelet madre la relación de ½ entre escalas es sólo aproximada, debido a efectos frontera en el filtrado). Entre los coeficientes de diferentes escalas del ejemplo anterior existirá una relación temporal: el segundo coeficiente d_3, el tercer y cuarto coeficiente d_2, y los coeficientes d_1 5 a 8 se encuentran en la misma "zona de influencia" temporal , que corresponde a las muestras 9 a 16 de la secuencia original. La figura 3 ilustra esta relación temporal. Para corregir el desfase producido a las escalas más elevadas (bajas frecuencias) se puede realizar un desfase cuyo valor depende de la escala, tal como se muestra en la figura 5.

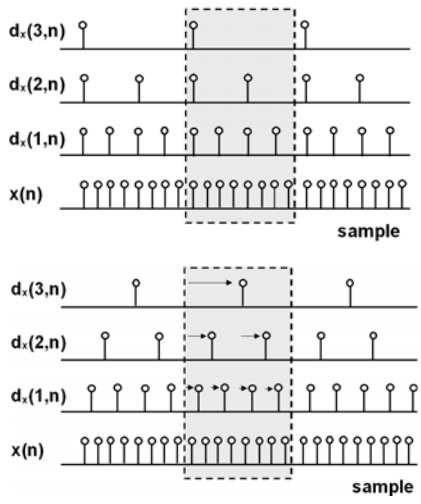


Figura 5: Arriba: Relación temporal entre las muestras de las escalas de la transformada wavelet discreta. Abajo: efecto del desfase dependiente de la escala.

Tras la fase de detección de los candidatos a punto de cambio en cada escala, se hace necesario realizar un proceso de agrupamiento y decisión (*clustering*) en función de su posición, con el objetivo de eliminar los puntos de cambio espurios creados por pequeñas no-estacionaridades de la varianza en alguna escala. La transformada de Hough [18], utilizada en procesado de imagen para detectar líneas rectas, es un buen candidato, y con ella se han realizado las pruebas descritas en este artículo.

4.2 SWT-SIC

Otra manera de solucionar el problema de la corrección de fase de la DWT, y de paso aumentar la precisión de la estimación de la posición de cambio, es utilizar la transformada wavelet estacionaria, la cual nos proporciona el mismo número de coeficientes a todas las escalas. Por lo demás, el algoritmo SWT-SIC es idéntico al DWT-SIC.

4.3 Pruebas con trazas sintéticas

Ambos algoritmos (DWT-SIC, SWT-SIC) han sido validados con trazas sintéticas antes de ser aplicados a trazas reales de tráfico. Las pruebas fueron realizadas con segmentos de FGN, creando una traza cuyo parámetro de Hurst es constante a trozos. A continuación describiremos uno de los experimentos.

La traza se compone de 131072 muestras. La primera mitad de la traza presenta H=0.8, y es seguida por dos segmentos de 32768 muestras cada uno con H=0.9 y H=0.7, en este orden. Por tanto existen dos puntos de cambio localizados en las posiciones n=65536 y n=98304. La figura 6 muestra la representación temporal de la secuencia. Los tres segmentos de FGN se crearon con una media común de 1024 y una varianza unitaria. Nótese que el valor medio de la secuencia inicial no tiene ninguna influencia en la media de las secuencias de coeficientes obtenidos en cada escala (que será cero, tal como se ha razonado anteriormente). La wavelet utilizada en el experimento es Haar (db1 o Daubechies 1).

La figura 7 muestra el resultado del algoritmo DWT-SIC (con corrección de fase) con un análisis de 8 escalas y un nivel de significación de 0.1. En la figura se muestran los puntos de cambio detectados a cada escala. Los dos puntos de cambio del parámetro de Hurst se pueden identificar claramente en las escalas bajas (1-3). A escalas superiores, uno de los puntos desaparece pero el otro se sigue detectando sin problemas. La razón de la desaparición del primer punto es la falta de precisión de la DWT a altas escalas, que provoca un efecto de suavizado que hace desaparecer un cambio de varianza que, al ser de carácter débil (dada la proximidad de los valores de H, 0.8 y 0.9) pasa desapercibido. El segundo punto de cambio se mantiene gracias a la mayor diferencia entre los valores de H (pasa de 0.9 a 0.7), aunque si hiciéramos una descomposición a escalas mayores acabaríamos perdiéndolo igualmente. Un fenómeno destacable en la figura 7 es la detección de puntos falsos de cambio (al principio de la secuencia, en la

escala 6). Estos espurios, en nuestros experimentos con trazas sintéticas aparecen típicamente al inicio y al final de la secuencia, son fácilmente detectables, ya que aparecen aislados en alguna escala, pero no en un número significativo de ellas. Otro fenómeno a comentar es la fluctuación de los puntos de cambio alineados alrededor de la posición 98304. Dicha fluctuación es provocada por la falta de precisión de la DWT a escalas altas, y aunque el desfase introducido por el algoritmo hace que su valor disminuya, no puede ser eliminado en su totalidad. Finalmente destacamos la aparición de "puntos ciegos" como el de la cuarta escala para el segundo punto de cambio.

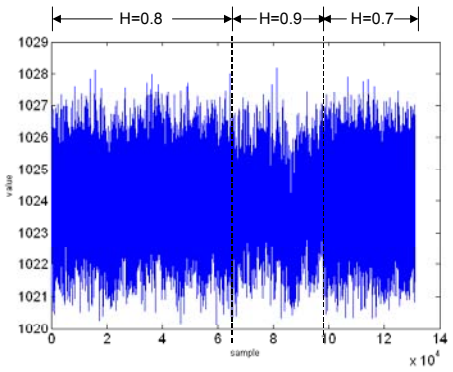


Figura 6: Traza sintética utilizada en la prueba, compuesta por tres segmentos de FGN.

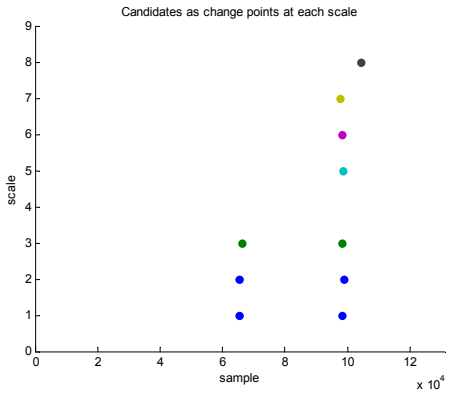


Figura 7: Resultado de DWT-SIC sobre la traza sintética. Análisis a 8 escalas, significación = 0.1.

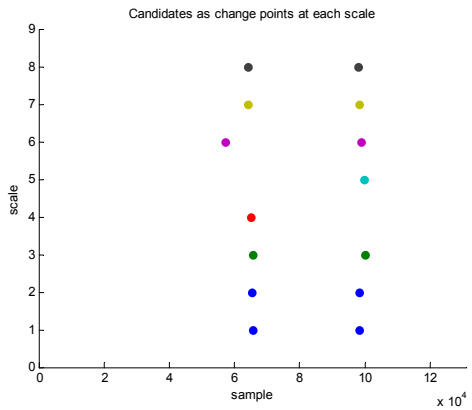


Figura 8: Resultado de SWT-SIC sobre la traza sintética. Análisis a 8 escalas, significación = 0.0001.

Con el objetivo de comparar el rendimiento de los dos algoritmos, en la figura 8 se muestra el resultado de aplicar el algoritmo SWT-SIC sobre la misma traza sintética. Ambos puntos de cambio son detectados, se repite el fenómeno de los puntos ciegos (en la escala 4 para el segundo punto, y en la escala 5 para el primero). Se observa un desplazamiento en la detección del primer punto de cambio en la 6ª escala. Esto es probablemente debido a que en dicha escala las varianzas de las dos subsecuencias alrededor del punto de transición todavía son demasiado cercanas (es decir, que el efecto de punto ciego de la escala 5 hace que la estimación sea "ruidosa").

A falta de un estudio en profundidad (que incluiría el efecto de la familia wavelet y del nivel de significación), podemos decir que en general la estimación con SWT-SIC ofrece mejores resultados que DWT-SIC en cuanto a la detección de puntos de cambio, pese a que suele ser más ruidosa. Por otro lado el algoritmo SWT-SIC requiere de mayor potencia computacional que el DWT-SIC, ya que al no producirse un diezmado en las escalas altas disponemos de muchas más muestras que con la DWT-SIC.

4.4 Pruebas con trazas reales

Los algoritmos basados en SIC se han aplicado también a la traza pAug89 de Bellcore. Dicha traza siempre se ha estudiado como un todo, obteniendo un valor global, medio, de los parámetros fractales, y en particular del parámetro de Hurst. Se ha documentado exhaustivamente que la traza presenta un valor de H alrededor de 0.8. Con la traza agregada a slots de 10 ms, el estimador LogScale devuelve un valor de H=0.803 (0.792, 0.813). Al aplicar los algoritmos DWT-SIC y SWT-SIC a la traza[3] obtenemos los resultados representados en las figuras 9 y 10, y en la Tabla 1.

Tabla 1 – Resultados del algoritmo DWT-SIC aplicado a la traza Bellcore pAug89 a 10 ms.

Segmento (muestras)	Parámetro de Hurst	Escalas
1-37938	0.822	j_1=3, j_2=9
37939-80546	0.790	j_1=3, j_2=8
80547-98297	0.836	j_1=3, j_2=12
98298-150905	0.810	j_1=4, j_2=14
150906-168858	0.731	j_1=5, j_2=7
168859-191830	0.825	j_1=3, j_2=13
191831-246958	0.737	j_1=2, j_2=6
246959-258227	0.688	j_1=2, j_2=12
258228-262144	0.804	j_1=4, j_2=8

[3] Debido a limitaciones de implementación, que por ahora sólo permite el análisis de trazas cuyas longitudes son potencias de 2, nuestro algoritmo se ha aplicado únicamente a las primeras 262144 muestras.

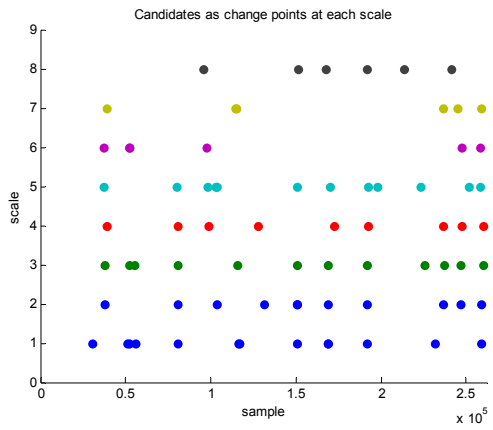


Figura 9: Resultado de DWT-SIC sobre la traza Bellcore pAug89 a 10 ms. Análisis a 8 escalas, significación = 0.1.

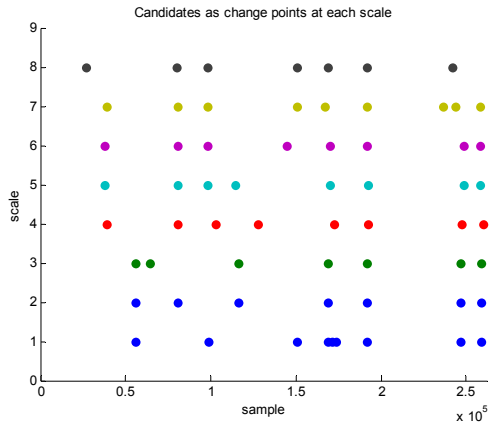


Figura 10: Resultado de SWT-SIC sobre la traza Bellcore pAug89 a 10 ms. Análisis a 8 escalas, significación = 0.0001.

En las figuras 9 se pueden observar claramente varios puntos de cambio alineados en un número significativo de escalas en las posiciones aproximadas n=38000, 80500, 98000, 151000, 169000, 192000, 247000, y 258000. Los resultados del proceso de clustering y decisión se resumen en la Tabla 1. Existen otros puntos de cambio, pero se han descartado por no estar alineados o por aparecer en un número de escalas inferior a 3.

El "rango dinámico" de los valores de H observados en cada segmento no es en absoluto despreciable: H varía entre 0.688 y 0.836. La traza parece mantenerse cerca de 0.8 durante las primeras 150000 muestras, para posteriormente oscilar de manera notable. Los valores son coherentes con el "valor medio" o global de 0.803. Los resultados coinciden en gran medida con estudios realizados previamente con la ICSS, el otro criterio de detección de cambios de varianza, que hemos estudiado en [9].

4.5 Hacia el análisis en tiempo real

Una vez comprobada la posibilidad de detectar los puntos de cambio, hemos realizado algunos pasos hacia un nuevo objetivo: el análisis y detección de transiciones en tiempo real. Obtener un algoritmo capaz de realizar el análisis a medida que entran las muestras de tráfico sería de indudable valor para los sistemas de monitorización y ciertos mecanismos que pueden reaccionar a cambios en las características fractales del tráfico.

La transformada DWT es fácilmente implementable en tiempo real. La SWT también, aunque su carácter redundante dificulta su escalado a muchas escalas de análisis, pero no es imposible. El punto problemático es la implementación en modo progresivo del algoritmo de detección de cambios de varianza; es decir, una versión en tiempo real del criterio SIC. La dificultad proviene del uso de la técnica de segmentación binaria, puesto que la cantidad de operaciones depende directamente del número de transiciones existentes, que es desconocido a priori.

Nuestros primeros pasos se han dirigido hacia una versión progresiva basada en ventana. Para determinar si era plausible hemos realizado diversos experimentos dirigidos a determinar si la estimación enventanada devuelve resultados coherentes; por coherencia entendemos, por ejemplo, que una transición detectada en una posición de ventana sigue siendo detectada en posiciones posteriores de la ventana. También hemos realizado experimentos para determinar la velocidad con que cada algoritmo es capaz de actualizar la estimación de las transiciones. Si estamos analizando una transición en la posición n, nuestros algoritmos no la detectan inmediatamente, sino que requieren algunas muestras adicionales. En [13,14,15] se pueden encontrar más detalles sobre la implementación en modo progresivo.

5 Conclusiones y líneas futuras

Este artículo describe el Criterio de Información de Schwarz y su aplicación al estudio de las variaciones del parámetro de escalado de tráfico con dependencia a largo plazo, usando una versión adaptada del estimador LogScale de Abry y Veitch. Se han presentado dos algoritmos, DWT-SIC y SWT-SIC, que presentan características diferentes: el basado en la SWT es más preciso en la localización de los puntos de transición, y puede proporcionar actualizaciones más rápidas cuando se usa en modo progresivo, mientras que el algoritmo DWT-sic es más rápido y supone una carga computacional menor. También se han presentado algunos resultados iniciales sobre la implementación de los algoritmos en modo progresivo, con el objetivo de desarrollar un analizador capaz de funcionar en tiempo real.

Se han presentado resultados de la validación de los algoritmos con trazas sintéticas, así como del análisis de una traza de tráfico real. Un resultado a destacar es la variación, nada despreciable, del parámetro de Hurst de la traza Bellcore, que hasta el día de hoy se había asumido constante. Puesto que el grado de fractalidad del tráfico tiene importantes implicaciones en el rendimiento de la red (anchos de banda efectivos, retardos en cola y ocupación de búfer

superiores a los predichos por los modelos markovianos y poissonianos), consideramos que es de gran interés caracterizar la evolución temporal de los parámetros LRD del tráfico, tal como permiten los algoritmos presentados.

Quedan muchos aspectos abiertos para estudios posteriores, entre los que destacan el cálculo de los intervalos de confianza de la estimación de las posición de las transiciones, la evaluación del coste computacional de los algoritmos, el desarrollo de una versión on-line que permita un análisis en tiempo real (o, al menos, progresiva) del tráfico, y una investigación profunda de la influencia de la familia wavelet usada en la estimación y del grado de significación, así como su significado en un entorno de medida de tráfico de red. También es de interés integrar en nuestro algoritmo los resultados recientemente publicados en [19], donde se demuestra que los coeficientes wavelet de las trazas reales de tráfico están distribuidos según una distribución normal generalizada; en estos momentos estamos desarrollando la adaptación de SIC (que asume una distribución normal) para extenderla a la normal generalizada. Finalmente, sería interesante extender los algoritmos a los modelos multifractales.

Nuestros esfuerzos tienen como objetivo desarrollar un estimador progresivo, en tiempo real, capaz de realizar una segmentación simultáneo en tiempo y frecuencia. Algunos de los algoritmos presentados han sido implementados en un analizador/generador de tráfico Gigabit Ethernet descrito en [16].

Referencias

[1] W. Leland et al, “On the self-similar nature of Ethernet Traffic”, IEEE/ACM Transactions on Networking, vol 2, pp 1-15, Feb 1994.

[2] K. Park, W. Willinger (eds), *Self-similar traffic and network performance*, Wiley, 2000.

[3] M. Roughan, D. Veitch, P. Abry, “On-line estimation of the parameters of long-range dependence”, Proceedings of IEEE Globecom 98, Nov. 1998, pp 3716 - 3721 vol.6

[4] D. Veitch and P. Abry, “A statistical test for the time constancy of scaling exponents”, IEEE Transactons on Signal Processing, vol. 49, no. 10, pp. 2325-2334, Oct. 2001.

[5] G. He, Y. Gao, J.C. Hou, K. Park, “A case for exploiting self-similarity of network traffic in TCP congestion control”, Proceedings of the 10th IEEE International Conference on Network Protocols, pp 34-43, November 2002.

[6] Y. Ouyang, L. Yeh, “Predictive bandwidth control for MPEG video: a wavelet approach for self-similar parameters estimation”, Proceedings of ICC 2001, pp 1551-1555, June 2001.

[7] X. Yu, I. Thng, Y. Jiang, “Measurement-based effective bandwidth estimation for long range dependent traffic”, Proceedings of IEEE Region 10 International Conference on Electrical and Electronic Technology, pp.359-365, Aug 2001.

[8] D. Veitch, P. Abry, “A wavelet-based joint estimator of the parameters of long-range dependence”, IEEE Transactions on Information Theory, vol 45, nº 3, April 1999.

[9] D. Rincón, S. Sallent, “Characterizing Fractal Traffic with Redundant Wavelet-based Transforms”, Proceedings of the First Workshop “New Trends in Modelling, Quantitative Methods and Measurements” - EuroNGI - Design and Engineering of the Next Generation Internet, pp.361-371, June 2004.

[10] J. Chen, K. Gupta, “Testing and Locating Variance Change points with Application to Stock Prices”, Journal of the American Statistical Assoc. vol 92, nº 438. June 1997.

[11] H. Akaike, “Information Theory and an Extension of the Maximum Likelihood Principle”, in Proceedings of the 2nd International Symposium of Information Theory, Budapest, pp267-281, 1973.

[12] G. Schwarz, “Estimating the dimension of a model”, The Annals of Statistics, 6, 461-464, 1978.

[13] D. Rincón, S. Sallent, “On-line Segmentation of Non-stationary Fractal Network Traffic with Wavelet Transforms and Log-likelihood-based Statistics”, LNCS 3375, pp. 110-123.

[14] D. Rincón, M. De Andrade, S. Sallent, “Towards On-line Segmentation of Fractal Network Traffic”, Infocom Student Workshop 2005.

[15] D. Rincón, S. Sallent, “Segmentation of Fractal Network Traffic with Wavelets and Log-likelihood Statistics”, IEEE ICC 2005, pp 1-5 (CD)

[16] D. Rincón et al, “Synthesis and analysis of fractal LAN traffic at high speeds”, Proceedings of IEEE LANMAN 04, pp.259-264, April 2004.

[17] V. Paxson, “Fast approximation of self-similar traffic”, Technical Report LBL-36750, Lawrence Berkeley National Laboratory, 1995.

[18] J.C.Russ, *The Image Processing Handbook (3rd edition)*, CRC Press, October 1998.

[19] P. Glomb, “Analysis of fGn and HTTP Requests Traces Using Localized Multiscale H Parameter Estimation”, IEEE SAINT 2005, pp 288-291, January 2005

Análisis de algoritmos de asignación de recursos a dos flujos de tráfico

Vicent Pla Boscà Vicente Casares Giner Jorge Martínez Bauset
Departamento de Communicaciones, ETSIT
Universidad Politécnica de Valencia (UPV)
Camí de Vera s/n, 46022 Valencia
Teléfono: 963879733, Fax: 963877309
(vpla,vcasares,jmartinez)@dcom.upv.es

Abstract *We study a family of admission control algorithms for a system with two arrival streams with different priority, which can be applied to handover prioritization in mobile cellular networks. These algorithms are based on the division of available servers into two types: non-reserved channels, that can be accessed by both types of arrivals; and reserved channels, that can only be accessed by high-priority arrivals (handovers). The parameter that specifies the amount of servers of each type varies continuously between zero and the total amount of available servers. In the queuing model of the system both traffic streams are treated according to a queuing model with two finite waiting lines. Besides, waiting customers may abandon their queues due to impatience. The queuing model is analyzed using a matrix-geometric approach. This paper is a compilation and an extension of previous studies appeared in the literature.*

1. Introducción

En este artículo se estudia una familia de algoritmos de control de admisión para sistemas con dos tipos de tráfico de distinta prioridad. Aunque las aplicaciones potenciales de estos algoritmos en el campo de las telecomunicaciones es muy diverso, la aplicación que ha motivado nuestro estudio es el control de admisión en redes móviles celulares. En este tipo de redes es necesario dar una prioridad mayor a las peticiones de traspaso —en lo sucesivo utilizaremos el término *handover*— que a la solicitudes de establecimiento de una nueva sesión, ya que, desde la perspectiva del usuario, la finalización forzosa de una sesión debido a un handover fallido resulta más molesta que el hecho de no poder establecer una sesión nueva.

El funcionamiento de los algoritmos estudiados se basa en separar los canales disponibles en dos tipos: los que pueden ser utilizados por cualquier tipo de petición y los reservados, que sólo pueden ser utilizados por las peticiones de alta prioridad (handovers). Ambos flujos de tráfico, las llamadas nuevas y las peticiones de handover, son tratados según un modelo de espera con cola finita, contemplándose la posibilidad de que la petición abandone la cola por impaciencia. En todos los algoritmos el parámetro que establece la cantidad de canales reservados puede variar de forma continua entre cero y el total de canales disponibles. Para la evaluación de prestaciones de los algoritmos se desarrolla un modelo analítico que utiliza una metodología geométrico-matricial [1, 2].

La especificación y el análisis de estos algoritmos supone una compilación y una extensión de varios estudios y propuestas que han aparecido en la literatura especializada. El algoritmo denominado *guard channel* (GC) —entre muchas otras formas— fue introducido como una técnica de CAC en redes celulares a mediados de los ochenta [3, 4]. En [5] se introduce la técnica *fraccional guard channel* (FGC) que es una generalización del GC en la que el número de canales reservados puede ser fraccionario. Schehrer [6] y, posteriormente, McMillan [7] proponen y analizan una extensión del GC que incorpora un ciclo de histéresis que gobierna la cantidad de recursos reservados. Además, en [7] se considera la posibilidad de espera y abandono por impaciencia para los dos tipos de llegada. Casares y Holtzman [8], y Casares [9] estudian toda una familia de extensiones al mecanismo GC en el contexto de un sistema de *trunking*. Del análisis de [8, 9] destaca el hecho de que se consideran tasas de servicio distintas para los dos tipos de llegada; en nuestro análisis, y en el resto de trabajos que aquí citamos, se supone que el tiempo medio de ocupación de los recursos en una célula es el mismo independientemente de si se trata de una llamada nueva o de una llamada que se traspasó desde otra célula. Los algoritmos que consideramos aquí son un subconjunto de los estudiados por Casares, sin embargo, los que aquí se analizan son más generales por cuanto incorporan la posibilidad de espera para las peticiones de alta prioridad, el parámetro de configuración de cada algoritmo puede ser fraccionario y en ambas colas se considera la posibilidad de abandono por impaciencia. En un trabajo anterior [10] ya habíamos considerado estos mecanismos en los que además se incorporaba un ciclo de histéresis como el

de [6, 7]. No obstante, en [10] no existe abandono para las llamadas nuevas que esperan y el parámetro de configuración de los algoritmos ha de ser entero. Kulavaratharasah y Aghvami [11] estudian mediante simulación dos de los mecanismos que aquí se consideran más un tercero que es una combinación probabilística de los dos primeros. Sin embargo, en el estudio de [11] el número de canales reservados es entero y no hay espera para ninguno de los dos tipos de petición. Daley y Servi [12, 13] introducen la reserva fraccionaria en los dos mismos mecanismos que en [11], sin embargo, aunque al final de [13] se sugiere la posibilidad de considerar la espera y el abandono por impaciencia, estos aspectos no se incluyen en los resultados del artículo.

El resto de este artículo esta estructurado del siguiente modo. En la sección 2 se especifica el comportamiento de los distintos algoritmos. Las secciones 3 y 4 se dedican a describir el modelo markoviano de un sistema en el que se aplican los algoritmos estudiados y se detallan los elementos principales para su análisis. La sección 5 muestra algunos resultados numéricos correspondientes a la evaluación de prestaciones de los algoritmos. Finalmente, en la sección 6 se resume el artículo.

2. Descripción de los algoritmos

Para la especificación de los algoritmos partimos de una descripción alternativa a la habitual del mecanismo FGC [5] en el que, además, se considera la posibilidad de espera para ambos tipos de peticiones. A partir de esta descripción alternativa el resto de mecanismos aparecen como una extensión natural del primero.

Los canales se reparten en tres grupos: grupo primario, grupo secundario y canal parcialmente reservado (grupo de un único canal). La idea general en todos los algoritmos es que los canales del grupo primario pueden ser asignados a cualquier tipo de petición —llamada nueva o handover— los del grupo secundario se reservan para las peticiones de mayor prioridad —handover— y el canal parcialmente reservado puede asignarse a una petición de llamada nueva sólo a veces —con cierta probabilidad—. C representa el número total de canales y el parámetro t ($0 < t < C$) establece el número de canales de los grupos primario y secundario, así como la probabilidad con la que el canal parcialmente reservado puede utilizarse para las llamadas nuevas. Así, si m representa el número de canales en el grupo primario, n en el grupo secundario y f la probabilidad de que una llamada nueva acceda al canal parcialmente reservado, se tiene que

$$m = \lfloor t \rfloor \qquad f = t - m \quad \text{y} \quad n = C - (m + 1).$$

De este modo, el número de canales no reservados es, en media, $m + f \cdot 1 = t$. Por tanto, mediante este mecanismo probabilístico se puede reservar una cantidad no entera de canales lo cual permite un ajuste más fino del control de admisión.

Para cada tipo de petición existe una cola de espera para las peticiones que no pueden ser atendidas en el momento de su llegada. Ambas colas son de capacidad finita y las llegadas a una cola llena se pierden. En el modelo se considera que las peticiones —ambos tipos— tienen una paciencia limitada por lo que pueden abandonar la cola antes de recibir servicio si el tiempo de espera excede un determinado límite.

En cada algoritmo se tienen tres tipos de evento que disparan la realización de una acción por parte del sistema de asignación de recursos. Los eventos posibles son: la llegada de una petición de handover, la llegada de una llamada nueva y la liberación de un canal. A continuación se describe para cada algoritmo la acción o acciones a realizar por el sistema en cada evento.

FGC (*Fractional Guard Channel*)

Llegada de una petición de handover: Intentar la asignación de un canal o, alternativamente, de una posición en la cola, siguiendo la secuencia: canal del grupo primario, canal parcialmente reservado, canal del grupo secundario, posición en la cola de handovers. Si ninguna de estas asignaciones es posible se rechaza la petición.

Llegada de una llamada nueva: Intentar la asignación de un canal o, alternativamente, de una posición en la cola, siguiendo la secuencia: canal del grupo primario; con probabilidad f, canal parcialmente reservado; Posición en la cola de llamadas nuevas. Si ninguna de estas asignaciones es posible se rechaza la petición.

Liberación de un canal: Si hay alguna petición de handover esperando en la cola, asignar el canal a la primera de ellas. Si no, reetiquetar los canales[1] e intentar la asignación del canal libre a la primera petición de la cola de llamadas nuevas siguiendo el mismo criterio que cuando llega una llamada nueva.

F-HOPSWR *(Fractional-Handovers Overflow from Primary to Secondary With Rearrangement)*

Llegada de una petición de handover o de una llamada nueva: Igual que en el algoritmo FGC.

Liberación de un canal: Realizar la primera de las acciones siguientes que sea posible: si hay alguna petición de handover esperando en la cola, asignar el canal a la primera de ellas; si hay alguna petición de llamada nueva en la cola, intentar la asignación del canal libre a la primera de ellas siguiendo el mismo criterio que cuando llega una llamada nueva; reetiquetar los canales.

F-HOPS *(Fractional-Handovers Overflow from Primary to Secondary)*

Llegada de una petición de handover o de una llamada nueva: Igual que en el algoritmo FGC.

[1] Se comienza por los canales ocupados y se les asignan las etiquetas en el orden siguiente: grupo primario, canal parcialmente reservado, grupo secundario. Las etiquetas sobrantes se asignan a los canales desocupados. De este modo primero se ocupan todos los canales del grupo primario, después el canal parcialmente reservado y, por último, los canales del grupo secundario.

Liberación de un canal: Realizar la primera de las acciones siguientes que sea posible: si hay alguna petición de handover esperando en la cola, asignar el canal a la primera de ellas; si hay alguna petición de llamada nueva en la cola, intentar la asignación del canal libre a la primera de ellas siguiendo el mismo criterio que cuando llega una llamada nueva.

F-HOSP *(Fractional-Handovers Overflow from Secondary to Primary)*

Llegada de una petición de handover: Intentar la asignación de un canal o, alternativamente, de una posición en la cola, siguiendo la secuencia: canal del grupo secundario, canal parcialmente reservado, canal del grupo primario, posición en la cola de handovers. Si ninguna de estas asignaciones es posible se rechaza la petición.

Llegada de una llamada nueva: Igual que en el algoritmo FGC.

Liberación de un canal: Igual que en el algoritmo F-HOPS.

3. Descripción del modelo

A la célula llegan peticiones de llamadas nuevas y de handover con unas tasas λ_n y λ_h, respectivamente. Ambos flujos de llegada siguen un proceso de Poisson. Los tiempos listados a continuación se describen mediante variables aleatorias exponenciales cuyo parámetro se indica entre paréntesis: duración de una llamada (μ_c), tiempo de permanencia en una célula (μ_r), tiempo de permanencia en el área de handover (μ'_r) y tiempo máximo de espera en cola de una llamada nueva (η). Por tanto, las variables aleatorias siguientes también seguirán una distribución exponencial, con los parámetros indicados: tiempo de ocupación de recursos ($\mu = \mu_c + \mu_r$) y tiempo máximo de espera en cola de una petición de handover ($\gamma = \mu_c + \mu'_r$). El número de posiciones en la cola de llamadas nuevas es Q_n y en la de peticiones de handover Q_h. Anteriormente se han introducido los parámetros siguientes: C, que representa el número total de canales; t ($0 < t < C$), que establece el número de canales de los grupos primario y secundario, y la probabilidad con la que el canal parcialmente reservado puede utilizarse para las llamadas nuevas. Así, $m = \lfloor t \rfloor$ es el número de canales en el grupo primario, $n = C - (m + 1)$ en el grupo secundario y $f = t - m$ la probabilidad de que una llamada nueva acceda al canal parcialmente reservado.

Las prestaciones del sistema se cuantifican mediante las probabilidades de bloqueo y abandono, que se representa mediante: probabilidad de bloqueo (abandono) de una llamada nueva P_b^n (P_a^n), probabilidad de bloqueo (abandono) de una petición de handover P_b^h (P_a^h). Por tanto, la probabilidad de pérdida de una petición —porque es bloqueada o abandona por impaciencia— será $P^n = P_b^n + P_a^n$ para las llamadas nuevas y $P^h = P_b^h + P_a^h$ para las peticiones de handover.

La descripción del estado del sistema no es la misma para todos los algoritmos sino que utilizamos dos distintas: una para los algoritmos FGC y F-HOPSWR, y otra para los algoritmos F-HOPS y F-HOSP. En todos los casos la representación elegida da lugar a un modelo que es un *proceso cuasi de nacimiento y muerte* (QBD), finito y no homogéneo [1]. A continuación se describen las dos alternativas:

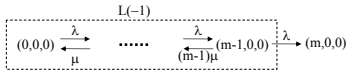
Algoritmos FGC y FHOPSWR El estado del sistema se representa mediante la terna de números enteros (k, i, j) : $0 \le k \le C,\ 0 \le i \le Q_n,\ 0 \le j \le Q_h$, donde k es el número de canales ocupados, i es el número de peticiones en la cola de llamadas nuevas y j es el número de peticiones en la cola de handovers. El *nivel* y la *fase* de cada estado no se corresponden directamente con ninguna de sus coordenadas —k, i o j— sino que se sigue el criterio siguiente. Si $L(l_0)$ representa el conjunto de estados del nivel l_0, entonces $L(-1) = \{(k, 0, 0) : k < m\}$ y $L(l_0) = \{(k, l_0, j) : k \ge m, 0 \le j \le Q_h\}$, $l_0 = 0, \ldots, Q_n$. Esto es, en el nivel -1 se agrupan todos los estados en los que el número de canales ocupados es inferior a m — todavía quedan canales libres en el grupo primario — y cuando el número de canales ocupados es igual o superior a m el nivel está determinado por el número de llamadas nuevas en la cola. En el nivel -1 el número de canales ocupados k representa la fase. En el resto de niveles la fase de un estado se corresponde con el número de canales ocupados que excede de m más el número de peticiones de handover en la cola, es decir, la fase del estado (k, l_0, j) es $k - m + j$.

En la figura 1 se representa en el diagrama de transiciones de los algoritmos FGC y F-HOPSWR. En cada nivel solo se han representado las transiciones internas y las de salida de ese nivel.

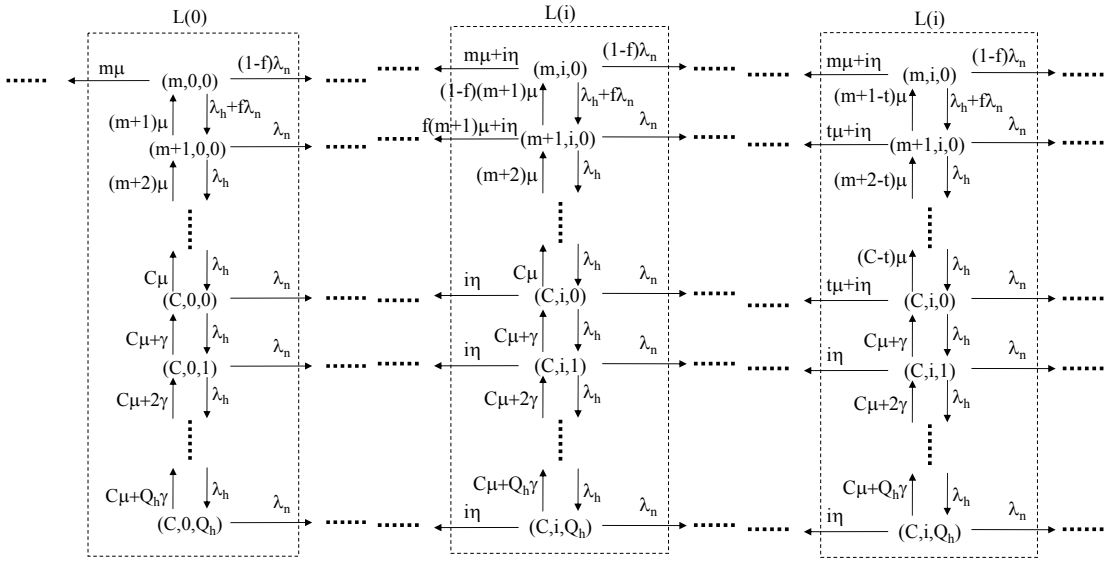
Algoritmos F-HOPS y F-HOSP El estado del sistema se representa mediante la quintupla de números enteros (i, j, r, k, l) : $0 \le i \le m;\ 0 \le j \le n;\ r = 0, 1;\ 0 \le k \le Q_n;\ 0 \le l \le Q_h$, donde i es el número de canales del grupo primario ocupados, j es el número de canales del grupo secundario ocupados, r indica si el canal parcialmente reservado está ocupado ($r = 1$) o no ($r = 0$), k es el número de peticiones en la cola de llamadas nuevas y l es el número de peticiones en la cola de handovers.

La agrupación de los estados en niveles es del siguiente modo $L(l_0) = \{(i, j, r, k, l) : i - m + k = l_0\}$, $l_0 = -m, \ldots, Q_n$. Dependiendo del número de estados (fases) podemos distinguir dos tipos de niveles, $l_0 = -m, \ldots, -1$:

$$L(l_0) = \{(l_0 + m, j, r, 0, 0) : 0 \le j \le n; r = 0, 1\},$$



(a) FGC y F-HOPSWR: nivel -1.



(b) FGC y F-HOPSWR: nivel 0. (c) FGC: nivel i. (d) F-HOPSWR: nivel i.

Figura 1: FGC y F-HOPSWR: diagrama de transiciones.

y, $l_0 = 0, \ldots, Q_n$:

$$L(l_0) = \{(m, j, r, k, l) : 0 \le j \le n; r = 0, 1; \\ 0 \le k \le Q_n; 0 \le l \le Q_h\}.$$

Al igual que antes, los niveles cuyo índice es negativo se corresponden con aquéllos en los que todavía quedan canales libres para acomodar la llegada de una llamada nueva. Los niveles del $-m$ al -1 tienen por tanto $2(n+1)$ fases, y los niveles del 0 al Q_n tienen $2(n+1)+Q_h$ fases. Así, el nivel del estado (i, j, r, k, l) es $i - m + k$ y, su fase $j + l + r(n+1)$.

En la figura 2 se representa en el diagrama de transiciones del algoritmo F-HOPS y en figura 3 el del algoritmo F-HOSP.

4. Análisis

Sea $\boldsymbol{\pi}$ el vector de las probabilidades estacionarias del proceso. Del mismo modo que con los estados, dividimos $\boldsymbol{\pi}$ en otros vectores más pequeños $\boldsymbol{\pi}^{(l_0)}$ correspondiendo cada uno de ellos a las probabilidades de estado de un nivel, por lo que el vector $\boldsymbol{\pi}^{(l_0)}$ tendrá tantas componentes como fases en el nivel l_0. El proceso que describe el comportamiento de cualquiera de los cuatro algoritmos es un proceso QBD, pues únicamente existen transiciones entre estados del mismo nivel o de dos niveles adyacentes y, en consecuencia, el generador infinitesimal del proceso tiene una estructura tridiagonal a bloques

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{A}_1^{(-J)} & \boldsymbol{A}_0^{(-J)} & & \\ \boldsymbol{A}_2^{(-J+1)} & \boldsymbol{A}_1^{(-J+1)} & \boldsymbol{A}_0^{(-J+1)} & \\ & & \ddots & \\ & & \boldsymbol{A}_2^{(Q_n)} & \boldsymbol{A}_1^{(Q_n)} \end{bmatrix} \tag{1}$$

siendo $J = 1$ para los algoritmos FGC y F-HOPSWR, y $J = m$ para los algoritmos F-HOPS y F-HOSP. En el apéndice A se puede encontrar el contenido de los de los distintos bloques de Q para los algoritmos FGC y F-HOPSWR. Debido a las limitaciones de espacio se omite el detalle de los valores de estas matrices para los otros dos algoritmos, F-HOPS y F-HOSP.

Las probabilidades de estado $\boldsymbol{\pi}$ se obtienen de la resolución del sistema de ecuaciones lineales $\boldsymbol{\pi Q} = \mathbf{0}^t, \boldsymbol{\pi e} = 1$. Donde $\mathbf{0}$ representa un vector columna de ceros y $\boldsymbol{e}$ un vector columna de unos.

Si $\boldsymbol{Q}$ es una matriz de dimensiones finitas, como es nuestro caso, este sistema en principio puede resolverse mediante cualquiera de los métodos estándar del álgebra lineal. Sin embargo, parece conveniente —sobre todo si el tamaño del sistema es grande— aprovechar la estructura y la naturaleza de $\boldsymbol{Q}$, que es un generador infinitesimal tridiagonal por bloques. Aquí hemos utilizado el algoritmo *Linear Level Reduction* [2, 14], que se aplica a la resolución de procesos QBD finitos y no homogéneos:

$$U \leftarrow \boldsymbol{A}_1^{(Q_n)}$$
$$R^{(Q_n)} \leftarrow \boldsymbol{A}_0^{(Q_n-1)} \left(-U\right)^{-1}$$

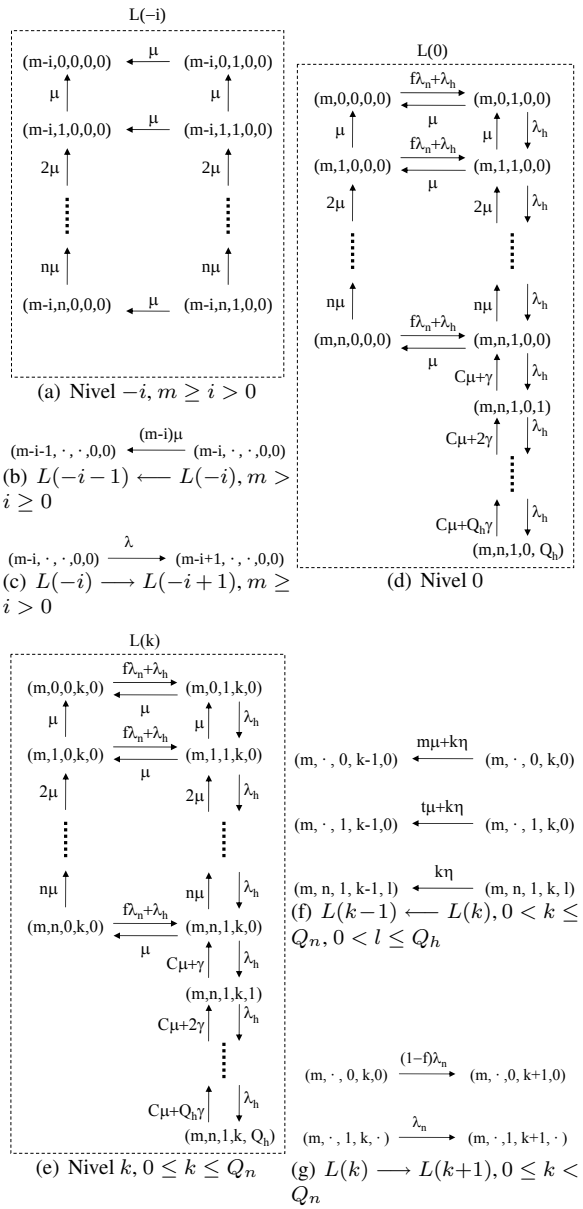


Figura 2: F-HOPS: diagrama de transiciones.

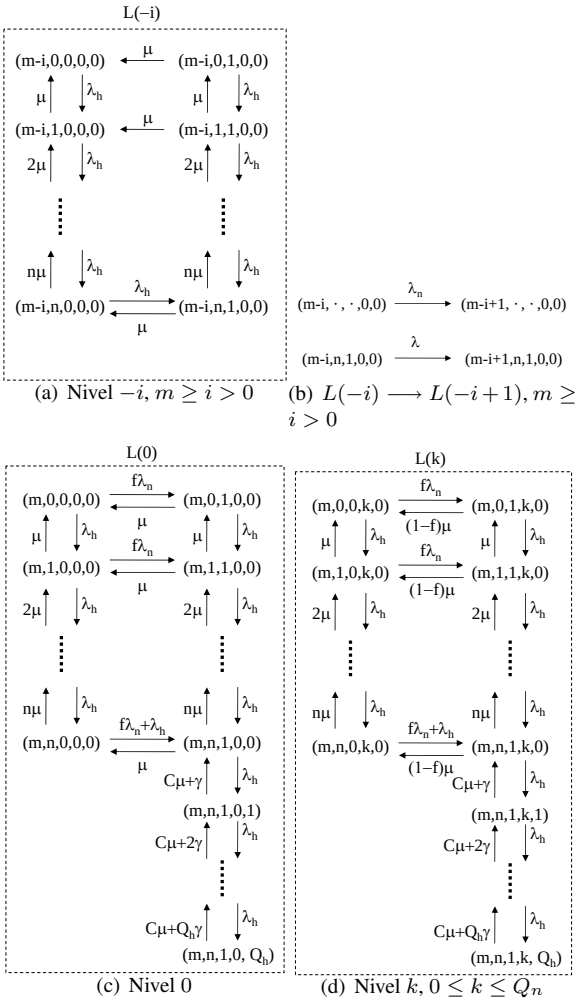


Figura 3: F-HOSP: diagrama de transiciones. Aquellas transiciones no recogidas en esta figura coinciden con las de la figura 2.

for $l = Q_n - 1, Q_n - 2, \ldots, 0, -J$ **do**
$U \leftarrow \boldsymbol{A}_1^{(l)} + R^{(l+1)} \boldsymbol{A}_2^{(l+1)}$
$R^{(l)} \leftarrow \boldsymbol{A}_0^{(l-1)} \left(-U\right)^{-1}$
end for

solve $\boldsymbol{\pi}^{(-J)}$ **from** $\{\boldsymbol{\pi}^{(-J)} U = \mathbf{0}^t; \boldsymbol{\pi}^{(-J)} \boldsymbol{e} = 1\}$
for $l = -J+1, \ldots, 0, \ldots, Q_n$ **do**
$\boldsymbol{\pi}^{(l)} = \boldsymbol{\pi}^{(l-1)} R^{(l)}$
end for

A partir de las probabilidades de estado los parámetros de medida de prestaciones se obtienen del siguiente modo. La probabilidad de bloqueo de las llamadas nuevas en los algoritmos FGC y F-HOPSWR se calcula mediante la expresión

$$P_b^n = \boldsymbol{\pi}^{(Q_n)} \Big[\overbrace{(1-f) \quad \cdots}^{n+1} \overbrace{\quad 1 \quad \cdots \quad 1}^{n+1+Q_n} \Big]^t,$$

y en los algoritmos F-HOPS y F-HOSP mediante la expresión

$$P_b^n = \boldsymbol{\pi}^{(Q_n)} \Big[\overbrace{(1-f) \quad \cdots \quad (1-f)}^{n+1} \overbrace{\quad 1 \quad \cdots \quad 1}^{n+1+Q_n} \Big]^t.$$

La probabilidad de abandono de las llamadas nuevas es, en todos los casos,

$$P_a^n = \frac{1}{\lambda_n} \sum_{r=1}^{Q_n} \boldsymbol{\pi}^{(r)} r \eta \boldsymbol{e}.$$

Para las peticiones de handover las probabilidades de bloqueo y abandono son, respectivamente,

$$P_b^h = \sum_{r=0}^{Q_n} \boldsymbol{\pi}^{(r)} \Big[0 \quad \cdots \quad 0 \quad 1 \Big]^t$$

y

$$P_a^h = \frac{1}{\lambda_n} \sum_{r=0}^{Q_n} \boldsymbol{\pi}^{(r)} \mu_r' \cdot \Big[0 \quad \cdots \quad 0 \quad 1 \quad 2 \quad \cdots \quad Q_h \Big]^t.$$

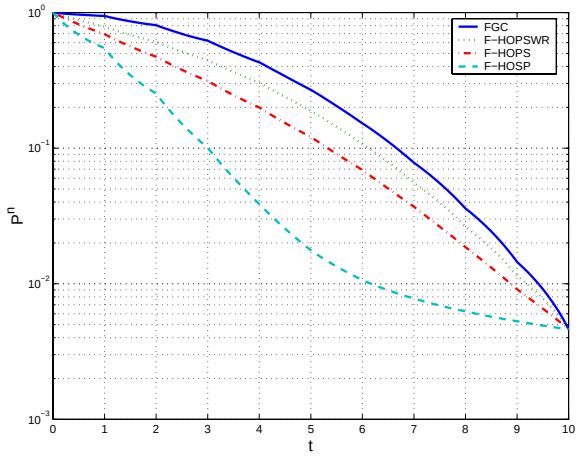


Figura 4: Probabilidad de fallo de llamada nueva P^n.

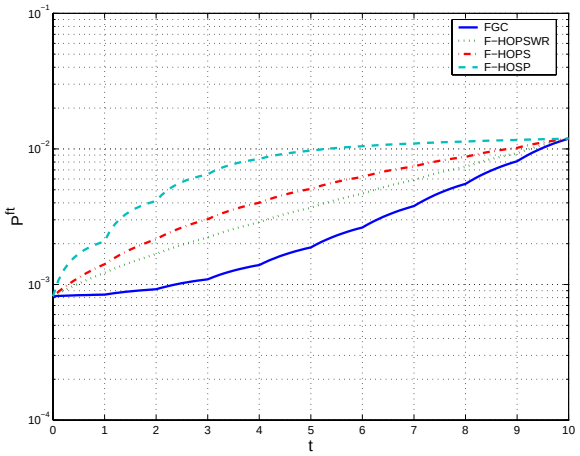
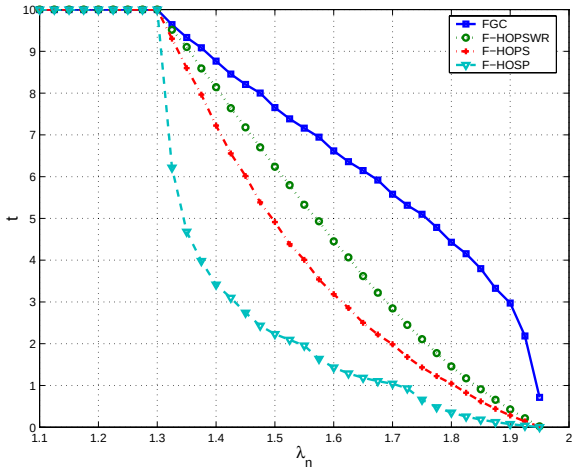
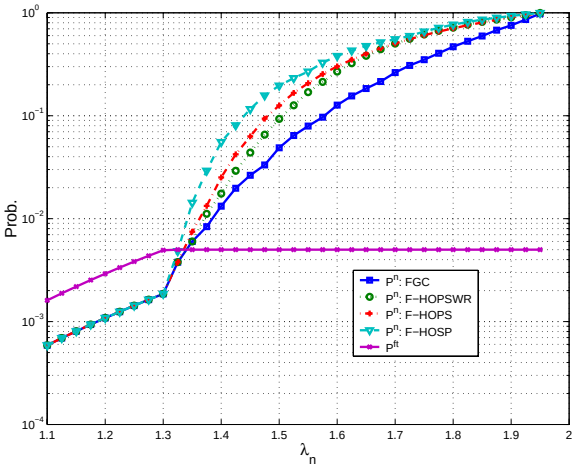


Figura 5: Probabilidad de terminación forzosa P^{ft}.



(a) valor de t



(b) valor de P^n y P^{ft}

Figura 6: Ajuste del parámetro t para que $P_{ft} \leq 0.005$.

5. Resultados numéricos

En los ejemplos numéricos mostrados a continuación se ha considerado la siguiente configuración: $\mu_r/\mu_c = 2$, $\mu_c + \mu_r = 1$ llamadas/s, $\lambda_n = 1.5$ llamadas/s, $\eta/\mu_c = \mu'_r/\mu_r = 10$, $C = 10$, $t = 8$, $Q_n = Q_h = 10$. En los distintos ejemplos se ha ido variando el valor de algunos parámetros de esta configuración básica. En las figuras 4 y 5 se muestra la variación, en función de t, de las probabilidades de pérdida de una petición nueva $P^n = P^n_b + P^n_a$, que puede deberse a la perdida inmediata por falta de espacio de almacenamiento (P^n_b) o al abandono por impaciencia (P^n_a), y de terminación forzosa $P^{ft} = \mu_r/\mu_c P^h/(1+\mu_r/\mu_c P^h)$, donde $P^h = P^h_b + P^h_a$ es la probabilidad de fallo de un handover, que también puede estar causado por el bloqueo debido a la falta de espacio de almacenamiento, o el abandono por impaciencia. Estas dos probabilidades varían de forma monótona y continua con el valor de t. De la observación de estas gráficas y de otras semejantes que no se muestran aquí, se desprende también que las funciones representadas aunque son continuas en todo el rango de variación de t pueden no ser derivables en los puntos en los que t toma un valor entero.

En las gráficas de la figura 7 se estudia el impacto del tamaño de las colas (Q_n y Q_h) sobre P^n y P^{ft}. El signo de este impacto es el esperado: un aumento de Q_n influye positivamente en P^n y negativamente en P^{ft}, mientras que el aumento de Q_h produce el efecto contrario. Sin embargo, lo más llamativo es que este efecto es prácticamente despreciable para tamaños de cola por encima de unas pocas unidades. Los resultados obtenidos con tasas de impaciencia menores ($\eta/\mu_c = \mu'_r/\mu_r = 2$) son cualitativamente semejantes: aunque inicialmente el impacto de aumentar el tamaño de las colas es mayor, éste se atenúa rápidamente y de nuevo deja de ser perceptible a partir de un tamaño de muy pocas unidades.

Finalmente, en la figura 6 se comparan las prestaciones de los diferentes algoritmos fijando un objetivo de QoS en función de la probabilidad de terminación forzosa $P^{ft} \leq 0.005$ y viendo cuál sería el valor de P^n en cada caso. La curvas de la figura 6(a) representan el valor máximo de t para el que se cumple el objetivo y en la gráfica de la figura 6(b) se representa P^n para el valor calculado de t. Según este criterio el algoritmo FGC es superior al resto, hecho este que ha sido demostrado formalmente para el caso particular en el que no hay colas ($Q_n = Q_h = 0$) en [13].

6. Conclusiones

En este artículo se ha estudiado una familia de algoritmos de control de admisión para sistemas con dos tipos de tráfico de distinta prioridad, como pueden ser las redes celulares. El funcionamiento de estos algoritmos se basa en separar los canales disponibles en dos tipos: los que pueden ser utilizados por cualquier tipo de petición y los reservados que sólo pueden ser utilizados por las peticiones de alta prioridad (handovers). Ambos flujos de tráfico son tratados según un modelo de espera con cola finita y, en el modelo, se contempla el abandono por impaciencia para ambos tipos de petición. En todos los algoritmos el parámetro que establece la cantidad de canales reservados puede variar de forma continua entre cero y el total de canales disponibles. El modelo analítico que se desarrolla en este trabajo utiliza una metodología geométrico-matricial. La especificación y el análisis de estos algoritmos supone una compilación y una extensión de varios estudios y propuestas que ha aparecido en la literatura especializada.

A. Bloques de Q

La tabla 1 detalla la dimensión y los valores de las matrices $\boldsymbol{A}_0^{(i)}$, $\boldsymbol{A}_1^{(i)}$, $\boldsymbol{A}_2^{(i)}$, ($i = -1, \ldots, Q_n$) correspondientes a los bloques del generador infinitesimal $\boldsymbol{Q}$ para los algoritmos FGC y F-HOPSWR.

Para la especificar el contenido estas matrices utilizamos la notación y convenciones siguientes. El operador $\mathrm{diag}\{\cdot\}$ (respectivamente: $\mathrm{diag}_1\{\cdot\}$ $\mathrm{diag}_{-1}\{\cdot\}$) devuelve una matriz cuya diagonal principal (respectivamente: diagonal de encima/debajo de la principal) es igual al vector argumento y el resto de entradas vale cero. El símbolo δ_i representa la delta de Kronecker, es decir $\delta_i = 1$ si $i = 0$ y 0 en cualquier otro caso, además $\delta_{i,j} = \delta_i \delta_j$.

Se utiliza $[\cdot]_{(i,j)}$ para referirse al elemento de la fila-i, columna-j de la matriz argumento. Los elementos de la diagonal de $\boldsymbol{A}_1^{(i)}$, que se representan mediante asteriscos, toman los valores necesarios para que las filas de $\boldsymbol{Q}$ sumen cero, es decir, $\boldsymbol{A}_1^{(-1)}\boldsymbol{e} + \boldsymbol{A}_0^{(-1)}\boldsymbol{e} = \boldsymbol{0}$ y $\boldsymbol{A}_2^{(i)}\boldsymbol{e} + \boldsymbol{A}_1^{(i)}\boldsymbol{e} + \boldsymbol{A}_0^{(i)}\boldsymbol{e} = \boldsymbol{0}$ ($i = 0, \ldots, Q_n$).

Agradecimientos

El presente trabajo ha sido financiado por la *Comisión Europea* (FEDER, 70 %) y el *Ministerio de Educación y Ciencia* (PGE, 30 %) a través de los proyectos TIC2003-08272 y TEC2004-06437-C05-01.

Referencias

[1] M. Neuts, *Matrix-geometric Solutions in Stochastic Models: An Algorithmic Approach*. The Johns Hopkins University Press, 1981.

[2] G. Latouche and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*. ASA-SIAM, 1999.

[3] E. C. Posner and R. Guérin, "Traffic policies in cellular radio that minimize blocking of handoff calls," in *Proceedings of ITC 11*, 1985.

[4] D. Hong and S. S. Rappaport, "Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures," *IEEE Transactions on Vehicular Technology*, vol. VT-35, no. 3, pp. 77–92, Aug. 1986, see also: CEAS Tech. Report No. 773, June 1, 1999, College of Engineering and Applied Sciences, State University of New York, Stony Brook, NY 11794, USA.

[5] R. Ramjee, R. Nagarajan, and D. Towsley, "On optimal call admission control in cellular networks," *Wireless Networks Journal (WINET)*, vol. 3, no. 1, pp. 29–41, 1997.

[6] R. Schehrer, "On a cut-off priority queueing system with hysteresis and unlimited waiting room," *Computer Networks and ISDN Systems*, vol. 20, pp. 45–56, 1990.

[7] D. McMillan, "Delay analysis of a cellular mobile priority queuing system," *IEEE/ACM Transactions on Networking*, vol. 3, no. 3, pp. 310–319, June 1995.

[8] V. Casares and J. Holtzman, "Dispatch versus interconnect traffic. a comparative analysis in a land mobile trunking system," in *Proceedings of the 46th VTC*, Apr. 1996, pp. 242–246, see for more details WINLAB TR-118, Rutgers University, NJ, May 1996.

[9] V. Casares-Giner, "Integration of dispatch and interconnect traffic in a land mobile trunking system. waiting time distributions," *Telecommunication Systems*, vol. 16, no. 3,4, pp. 539–554, 2001, previously presented at 4th INFORMS Telecommunications Conference, Boca Ratón (Florida) March 8–11, 1998.

[10] V. Pla and V. Casares, "Delay-loss analysis of channel assignment schemes in mobile cellular with handoff priority and hysteresis control," in *Proceedings of 14th ITC Specialist Seminar on Access Networks and Systems*, 2001, pp. 221–230.

[11] M. D. Kulavaratharasah and A. H. Aghvami, "Teletraffic performance evaluation of microcellular personal communication networks (PCN's) with prioritized handoff procedures," *IEEE Transactions on Vehicular Technology*, vol. 48, no. 1, pp. 137–152, Jan. 1999.

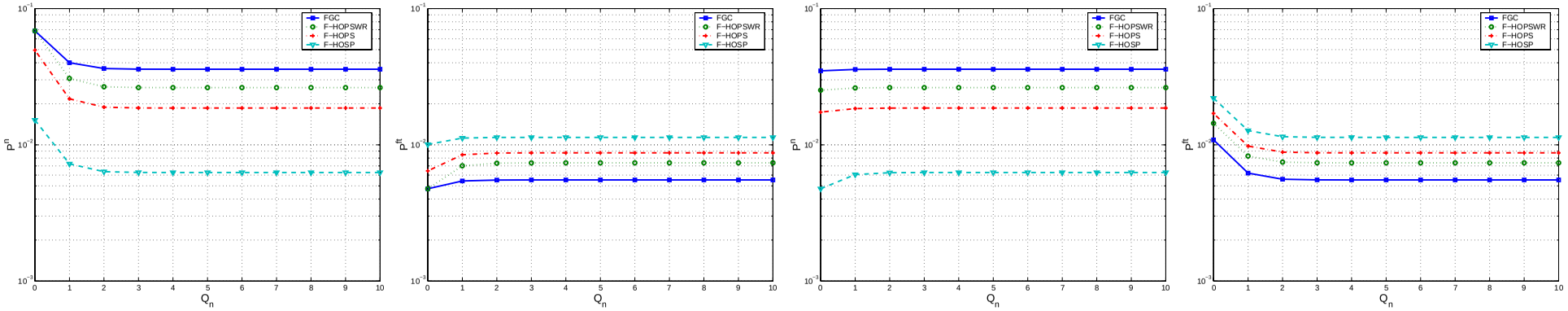
[12] D. J. Daley and L. D. Servi, "Loss probabilities of hand-in traffic under various protocols. I. models and algebraic results," *Telecommunication Systems*, vol. 19, no. 2, pp. 209–226, 2002.

[13] ——, "Loss probabilities of hand-in traffic under various protocols: II. model comparisons," *Performance Evaluation*, vol. 55, no. 3-4, pp. 231–249, Feb. 2004.

[14] D. Gaver, P. Jacobs, and G. Latouche, "Finite birth-and-death models in randomly changing environments," *Advances in Applied Probability*, vol. 16, pp. 715–731, 1984.

nivel (i)	dimensión	FGC	F-HOPSWR
		$\boldsymbol{A}_0^{(i)}$	
-1	$m \times (n+2+Q_h)$	$[\cdot]_{(i,j)} = \lambda\delta_{i-m,j-(n+2+Q_h)}$	$=$
$0,\ldots,Q_n-1$	$(n+2+Q_h)\times(n+2+Q_h)$	$\lambda_n \operatorname{diag}\{[1-f \quad 1 \quad \cdots \quad 1]\}$	$=$
		$\boldsymbol{A}_1^{(i)}$	
-1	$m \times m$	$\mu\operatorname{diag}_{-1}\{[1 \quad 2 \quad \cdots \quad m-1]\} + \lambda\operatorname{diag}_1\{\boldsymbol{e}\} + \operatorname{diag}\{[* \quad \cdots \quad *]\}$	$=$
0	$(n+2+Q_h)\times(n+2+Q_h)$	$\mu\operatorname{diag}_{-1}\{[(m+1)\mu \quad \cdots \quad C\mu \quad C\mu+\gamma \quad \cdots \quad C\mu + Q_h\gamma]\} + \operatorname{diag}\{[* \quad \cdots \quad *]\} + \lambda\operatorname{diag}_1\{[\lambda_h + f\lambda_n \quad \lambda_h \quad \cdots \quad \lambda_h]\}$	$=$
$1,\ldots,Q_n$	$(n+2+Q_h)\times(n+2+Q_h)$	$\mu\operatorname{diag}_{-1}\{[(1-f)(m+1)\mu \quad (m+2)\mu \quad \cdots \quad C\mu \quad C\mu+\gamma \quad \cdots \quad C\mu+Q_h\gamma]\} + \operatorname{diag}\{[* \quad \cdots \quad *]\} + \lambda\operatorname{diag}_1\{[\lambda_h + f\lambda_n \quad \lambda_h \quad \cdots \quad \lambda_h]\}$	$\mu\operatorname{diag}_{-1}\{[((m+1-t)\mu \quad \cdots \quad (C-t)\mu \quad C\mu+\gamma \quad \cdots \quad C\mu+Q_h\gamma]\} + \operatorname{diag}\{[* \quad \cdots \quad *]\} + \lambda\operatorname{diag}_1\{[\lambda_h + f\lambda_n \quad \lambda_h \quad \cdots \quad \lambda_h]\}$
		$\boldsymbol{A}_2^{(i)}$	
0	$(n+2+Q_h)\times m$	$[\cdot]_{(i,j)} = m\mu\delta_{i-1,j-m}$	$=$
$1,\ldots,Q_n$	$(n+2+Q_h)\times(n+2+Q_h)$	$i\eta\boldsymbol{I}_{n+2+Q_h} + \mu\operatorname{diag}\{[m \quad f(m+1) \quad 0 \quad \cdots \quad 0]\}$	$i\eta\boldsymbol{I}_{n+2+Q_h} + \mu\operatorname{diag}\{[m \quad \overbrace{t \quad \cdots \quad t}^{n+1} \quad 0 \quad \cdots \quad 0]\}$

Tabla 1: FGC: bloques del generador del generador infinitesimal $\boldsymbol{Q}$.



(a) Influencia del valor de Q_n en P^n. (b) Influencia del valor de Q_n en P^{ft}. (c) Influencia del valor de Q_h en P^n. (d) Influencia del valor de Q_h en P^{ft}.

Figura 7: Influencia del tamaño de las colas.

Algoritmo para el cálculo de topologías bi-conexas con restricciones de diámetro y su aplicación en el diseño de redes

K.-D. Hackbarth *), A. Menéndez *), C. Díaz**),J.A. Portilla **)
*) Universidad de Cantabria Dept. Ingeniería de Comunicación
**) Universidad de Alcalá Dept. Teoría de la Señal y Comunicaciones

e-mail klaus@tlmat.unican.es

Abstract. *This paper presents an algorithm which calculates a bi-connected graph with diameter limitation starting from a totally meshed graph. The problem is reduced into two separate subproblems: partition and augmentation. The first one reduces the initial set of edges considering weights assigned to themselves until fulfilling a threshold. If the solution of the first subproblem does not result a bi-connected graph, the algorithm inserts in the second part additional edges until the bi-connectivity of the graph is assured. The algorithm is applied to the topologic design of the routing layer for Next Generation Internet core network (NGI). The algorithm forms part of a general planning and dimensioning software tool. The objective of this software tool is to provide an optimal design of the logical network part (Routing layer) of NGI considering different types of traffic resulting from the subscribers connected to the NGI over corresponding access technologies.*

1 Introducción

Esta contribución describe un algoritmo para el cálculo de topologías de redes de comunicación bajo criterios de bi-conectividad y restricciones relativas al número de saltos máximo entre parejas de nodos expresado con el diámetro de un grafo.

El algoritmo se basa en dos principios fundamentales: reducción y expansión. De esta manera, en primer lugar se reduce una topología de red completamente mallada a un grafo conexo, no necesariamente bi-conexo, que cumpla la limitación del diámetro y después, se expande esta topología hasta el cumplimiento de la bi-conectividad. Tanto en la reducción como en la expansión se consideran umbrales relativos al peso de los enlaces. El peso asociado a cada enlace, se obtiene a través de una combinación de diferentes criterios, principalmente relacionados con su longitud y flujo, lo que permite la aplicación del algoritmo sobre una amplia variedad de problemas en el campo del diseño de topologías.

El algoritmo se aplica en concreto al diseño de la capa lógica de una red IP de nueva generación (NGI) donde la restricción en el diámetro es un parámetro importante. El algoritmo forma parte de una herramienta denominada TAROCA-IP para el diseño de una NGI. La aplicación principal de TAROCA-IP reside en estudios tecno-económicos ligados a los problemas de regulación que determinan las correspondientes Autoridades Nacionales de Regulación de Telecomunicación ANRT.

1.1 Definición del problema

El modelado de redes de comunicación, incluso de redes de otra naturaleza, utiliza con frecuencia modelos de la Teoría de Grafos, aplicando los algoritmos correspondientes [1]. La incursión de la Teoría de Grafos se debe fundamentalmente a la facilidad con la que muchos de los problemas que se plantean en el diseño de redes se trasladan a problemas de grafos, y así como las soluciones obtenidas se transforman con facilidad en soluciones reales.

Uno de los problemas que se plantea en el diseño de topologías para redes de comunicaciones reside en asegurar la conectividad entre todos los nodos en el caso de una avería o malfuncionamiento de uno de sus enlaces. Por otro lado, el principal objetivo en el diseño de una red de comunicaciones consiste en hallar una solución óptima en términos de sus costes de inversión (CAPEX) y operación (OPEX), lo que resulta normalmente una topología en forma de árbol. Una topología bi-conexa establece un compromiso entre la conectividad de la red y la capacidad de recuperación de la misma tras un fallo. En caso de que los costes de infraestructura dominen la estructura bi-conexa debe ser calculada tomando como punto de partida una topología de tipo spanning tree, ver [2].

Esta solución puede formar una base inicial para conseguir una topología bi-conexa mediante la introducción de enlaces adicionales. Sin embargo, la aplicación de este algoritmo en el campo de las redes de comunicaciones no resulta conveniente por dos motivos principales. En primer lugar, este problema no tiene en cuenta los costes reales compuestos por agregaciones de pesos de los enlaces entre los nodos; y en segundo lugar, no contempla la necesidad de mantener ciertos enlaces clave en la topología como bien pueden ser aquellos que unen las ciudades más pobladas de un país. Por otro lado, existen estudios y resultados que concluyen en el hecho de que es necesario limitar el máximo número de saltos en una

red de datos, lo que significa considerar un grafo con una restricción en el diámetro[1], ver [1].

Estas condiciones requieren un nuevo planteamiento del problema y una nueva aproximación para el diseño de la topología. Este diseño se aplica en dos niveles: en el nivel de enrutamiento/conmutación, donde las enlaces determinan las capacidades de señales eléctricas/ópticas entre los enrutadores/ conmutadores y en el nivel de transmisión, donde los enlaces determinan las infraestructuras de medios de transmisión, generalmente cables de fibra óptica. El primero se conoce también como capa lógica y el segundo como capa física. El diseño de la capa lógica parte normalmente de una red totalmente mallada y trata de reducir, en primer lugar, el número de enlaces a través de ciertos criterios de eliminación determinados por factores relativos a la demanda de trafico entre los nodos, la longitud de enlaces, el grado de conectividad y el número máximo de saltos en los caminos (problema de reducción).

De cualquier modo, tras la reducción la topología obtenida puede presentar una falta de caminos múltiples entre parejas de nodos, requeridos para las funciones de protección de la capa lógica. Con el objeto de construir una red más fiable, es necesario determinar una topología que proporcione k caminos disjuntos entre cualquier par de nodos; con al menos con k=2. Este número k se conoce como número de conectividad por enlaces o vértices[2] [1].

Para modelar el problema, supongamos un grafo sin dirección $\mathrm{G(V,E,}\omega)$, con pesos $\omega_i \in \omega$ para todos los arcos $e_i \in E$ dados. Con la aplicación del problema de reducción, el algoritmo determina un conjunto de arcos $E_o \subset E$, considerando en todo momento la limitación del diámetro. El objetivo del problema de partición y expansión reside en encontrar un sub-conjunto de arcos de expansión $EXP \subset E - E_0$ tales que $G(V, E_0 \cup EXP)$ sea un grafo al menos bi-conexo.

Este problema se muestra como un problema NP-hard [3]. En general, es computacionalmente costoso resolver un problema como éste de manera óptima para redes de tamaño real, con cientos de nodos, ya que las técnicas exactas requerirían demasiado tiempo y/o memoria. Por lo tanto, e consideran métodos heurísticos, tanto en el problema de reducción como en el de expansión, capaces de cuasi-optimizar el problema con dependencia polinómica respecto al número de nodos dorsales.

[1] El diámetro Δ en un grafo define el número máximo de saltos del camino de salto mínimo entre todas las parejas de nodos.

[2] En la terminología de la teoría de grafos, se aplica vértices para nodos y arcos para enlaces.

1.2 Bi-conectividad y algoritmos correspondientes

Según [5], un grafo G es k-conexo en arcos, si existe un subconjunto de k arcos con cuya eliminación resulta que el grafo G se convierte en no conexo; la k-conectividad por vértices se define de manera análoga. Para el caso particular de un grafo bi-conexo por arcos, normalmente definido como bi-conexo, será necesario eliminar al menos dos arcos para desconectar algunos vértices entre sí y convertirlo en no conexo. De este modo, si tenemos un grafo bi-conexo y con la condición de que se produzca un único fallo al tiempo, todos los nodos de la red serán alcanzables en todo momento. Para el problema de biconectividad, se considera que un puente e_p es un arco cuya eliminación del grafo añade una nueva componente no conexa al mismo. Para el problema bi-conexo en vértices se considera que un vértice v_a es un punto de articulación si la eliminación del mismo del conjunto de vértices que forman el grafo provoca la existencia de una componente no conexa añadida. Hay varias condiciones que expresan la bi-conectividad de un grafo, ver [6], donde la más importante resulta, según [7], que un grafo $\mathrm{G(V,E)}$ es k-conexo si para cada par de nodos x e y, existen al menos k caminos disjuntos que los unan.

Para el diseño de grafos bi-conexos, Frederickson, [8] propuso un algoritmo aproximativo para expandir un árbol a una topología bi-conexa y Khuller [9], un algoritmo similar con mejor comportamiento temporal. El algoritmo expuesto en este trabajo expande los anteriores incluyendo la restricción del diámetro y pesos para proporcionar prioridades a arcos y vértices, condiciones requeridas en el diseño de las redes como NGI.

2 Descripción del Algoritmo

El algoritmo aproximativo se compone de los siguientes pasos:

1. Extracción de un arco del conjunto inicial E.
2. Cómputo de los nuevos pesos de los arcos del conjunto E.
3. Construcción de un nuevo grafo G' formado un árbol compuesto por el conjunto de subgrafos conexos y los puentes que los unen.
4. Expansión del árbol G' mediante la introducción de nuevos arcos.
5. Retransformación de G' a su grafo original y determinación del conjunto de arcos de expansión.

Los pasos se agrupan en dos bloques diferenciados. El primero de ellos, genéricamente conocido como problema de reducción, está compuesto por los pasos 1 y 2. El segundo de los bloques se corresponde con los pasos 3 a 5 y se divide a su vez en otros dos sub-problemas: el primero, partición, se corresponde con el paso 3 y el segundo, expansión, abarca los pasos 4 a 5. El sub-problema de partición consiste en el cálculo de los puentes que unen los sub-grafos bi-conexos. La existencia de arcos puente se traduce en la existencia de cuellos de botella en la conectividad e implica la necesidad de insertar de nuevos arcos con objeto de lograr una topología bi-conexa. El sub-problema de expansión se encarga de definir los criterios bajo los cuales se insertan los nuevos arcos.

2.1 Problema de reducción

Sea E el conjunto de arcos de la topología inicial completamente mallada. El objetivo del problema de reducción consiste en la disminución del número de arcos del conjunto E obteniendo un subconjunto $E_o \subset E$. Esta reducción se realiza con la agregación de los pesos relativos a los arcos y considerando la limitación en el número de saltos.

La eliminación de cada arco produce un incremento en los flujos de algunos de los arcos restantes lo que significa también un aumento correspondiente de los valores de sus pesos. La reducción finaliza cuando los pesos en los arcos restantes son superiores o iguales a un umbral o bien se sobrepasa la limitación del diámetro. En el diseño de redes de telecomunicaciones, la función de peso puede estar compuesta por varios elementos como flujo de tráfico, número de usuarios en los vértices incidentes del arco, longitud geográfica del arco etc.

De esta manera, el algoritmo considera dos umbrales, un umbral de peso $\Omega_{\min}$ que determina si un enlace es o no candidato a ser eliminado, y un segundo umbral de diámetro, $\Delta_{\max}$, que se corresponde con el número máximo de saltos del camino de saltos mínimo entre los vértices.

El funcionamiento del algoritmo de reducción es el siguiente: Se recorren todos los arcos del conjunto E y para cada arco e_i, perteneciente a E, si el peso del arco, ω_i, es menor que el umbral $\Omega_{\min}$, el arco se convierte en candidato a ser eliminado. Entonces, el algoritmo se encarga de comprobar que todos los nodos afectados tras la eliminación del arco e_i puedan ser alcanzados a través de un camino con menor o igual número de saltos que el umbral $\Delta_{\max}$. Si ambas condiciones se cumplen entonces el arco e_i es eliminado del conjunto E y se actualiza la función de peso en los arcos restantes. En la Fig. 1 se expone el pseudo-código correspondiente.

```
for each e_i ∈ E
 if (ω_i < Ω_min)
    deactivate e_i
    if n_hops(e_i) < Δ_max
       remove e_i from E
       recalculate ω_i for e_i ∈ E
    else
       reactivate e_i
   end if
 end if
end for
```

Figura 1: Pseudo-código del Algoritmo de reducción

En la Fig.2, se muestra a través de un ejemplo el mecanismo de ejecución. Supongamos que el arco $e_4 = [2,4]$ es un candidato a ser eliminado ya que su peso es menor que el umbral establecido, ver fig. 2.a. En la fig.2.b, el algoritmo verifica que el número de saltos entre los vértices involucrados sea menor o igual que el umbral (supongamos un valor de tres saltos). Como se observa en el ejemplo, si se elimina el arco e_4, para ir de 2 a 4 habrá que pasar por el nodo 5 lo que significa un camino de dos saltos. De esta manera, se elimina el arco e_4 ya cumple las dos condiciones (función de peso y umbral de saltos).

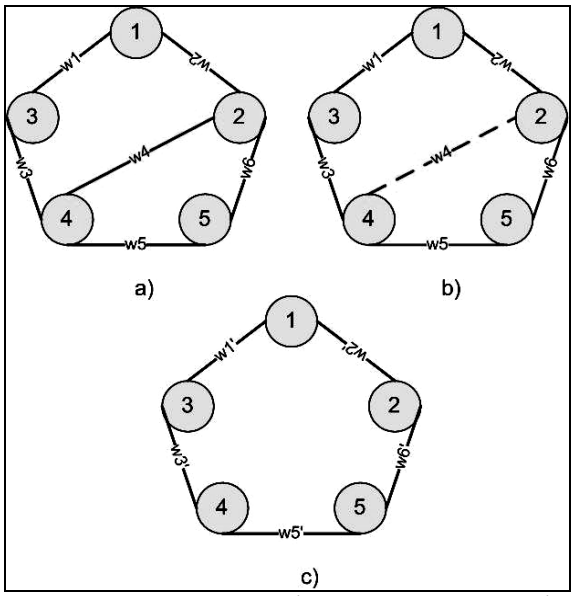


Figura 2: Ejemplo de aplicación del algoritmo de reducción

Cabe mencionar que puede darse el caso de que el problema de reducción obtenga una solución bi-conexa, en cuyo caso el algoritmo finalizaría en este punto Esta situación se produce principalmente cuando el valor del peso de los arcos iniciales E es ya alto o bien los umbrales $\Omega_{\min}$ o $\Delta_{\max}$ son bajos. En cualquier caso, la solución obtenida depende del orden en el que los arcos son analizados y por ello el algoritmo calcula una de las posibles soluciones existentes, el criterio para la ordenación de los arcos depende de la aplicación.

2.2 Problema de partición

El problema de partición se corresponde con la búsqueda de los arcos puente del grafo, donde un puente es un arco cuya eliminación descompone el grafo en dos partes no conexas. Para la búsqueda de los puentes se aplica un algoritmo DFS (Depth First Search), que proporciona un método de búsqueda a través del grafo, basado en un método de Tarjan, ver [10], [11]. Intuitivamente, la búsqueda comienza seleccionando algunos nodos del grafo como nodos raíz y se explora cada rama hasta donde sea posible. En la Fig.3 se muestra el pseudo-código del algoritmo de partición, incluyendo el algoritmo DFS.

Para cada arco e_i perteneciente al conjunto E_0, se desactiva el arco en primer lugar y se aplica sobre el grafo resultante, G, el algoritmo DFS. El arco e_i es un arco puente, y se almacena en el subconjunto E_1, si la aplicación de $DFS(G)$ concluye que alguno de los nodos pertenecientes al grafo G no es alcanzable a través sus arcos. A continuación, se reactiva de nuevo el arco e_i y se continúa evaluando el resto de arcos hasta que finaliza el algoritmo.

```
for each e_i ∈ E
    deactivate e_i
    DFS(G) if n_hops(e_i) < Δ_max
    if any node is unachievable
        remove e_i ∈ E_1
    end if
    reactivate e_i
end for
```

Figura 3: Algoritmo de partición

Identificados los puentes, se divide el grafo G en sub-grafos bi-conexos resultantes de la eliminación de los arcos puentes. Una vez identificados los sub-grafos, se forma un nuevo grafo, G', con "supervértices", correspondientes con cada uno de los sub-grafos bi-conexos, y con los arcos puentes que conectan los supervértices. A cada uno de estos supervértices v_i se le asigna un peso equivalente ψ_i cuyo valor se calcula en base a los pesos equivalentes de cada sub-grafo original. El cálculo de estos pesos equivalentes depende del criterio de optimización, definido particularmente para cada caso real. En la Fig. 4 se muestra un ejemplo en el que el algoritmo de partición encuentra tres arcos puente, lo que implica cuatro componentes bi-conexas. De esta manera, se forma un nuevo grafo con cuatro supervértices y tres arcos, que son precisamente los arcos puente.

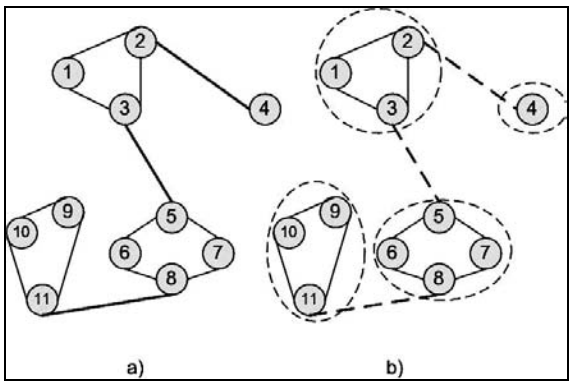


Figura 4: Funcionamiento del algoritmo de partición

2.3 Problema de expansión

Una vez que se han obtenido el grafo G' con sus supervértices y los arcos puentes, el cometido del algoritmo de expansión es introducir el número suficientes de arcos adicionales de manera que el grafo resultante G' sea bi-conexo. Como ya se mencionó en la sección anterior, el primer paso en el algoritmo de expansión es el cálculo de los pesos equivalentes relativos a cada sub-grafo bi-conexo. Con objeto de que la resolución del problema sea sencilla, se asume que cada uno de los subgrafos obtenidos a través del algoritmo de partición se convierte en un único vértice v_i con su correspondiente peso ψ_i. Estos pesos se utilizan como criterio para insertar arcos adicionales hasta que el grafo G' sea bi-conexo. En la Fig. 5 se muestra el resultado de la aplicación de los primeros pasos del algoritmo de expansión sobre el ejemplo de la Fig. 4. Así mismo, el pseudo-código de esta parte del algoritmo se expone en la Fig. 6.

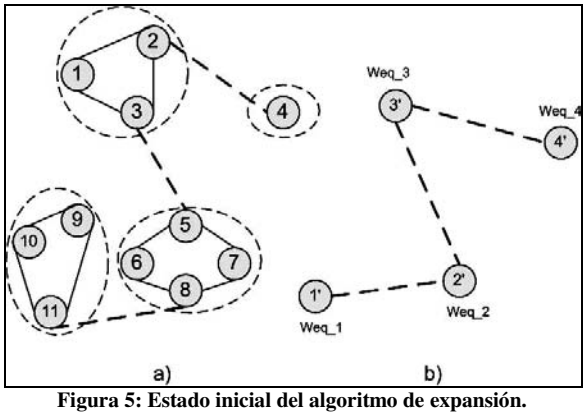


Figura 5: Estado inicial del algoritmo de expansión.

Como se puede observar, ver Fig. 6, el algoritmo recalcula los pesos equivalentes correspondientes a cada vértice y además recalcula el grafo G', formado por el nuevo conjunto de vértices. El siguiente paso es entonces la búsqueda de parejas de vértices entre los que insertan nuevos arcos. Se selecciona un vértice v_i de G' con grado[3] igual a uno y se busca el

[3] El grado de un vértice corresponde al numero de arcos que coinciden con el

vértice v_j de G' con máximo valor de peso ψ_i y con la condición de que no exista ya un arco entre dichos vértices v_i y v_j. Cuando finaliza el algoritmo de expansión, se transforma el grafo formado por los super-vértices en el grafo original.

```
recalculate  ω_i
recalculate  G'
for each  v_i ∈ G'  with  grade(v_i) = 1
    search  v_j ∈ G'  with  max(ψ_i)
end for
transform  G'  into  original  graph  G
```

Figura 6: Algoritmo de expansión

En el proceso de retransformación es necesario considerar que el algoritmo ha calculado los arcos entre parejas de supervértices, que se corresponden en realidad con el subconjunto de arcos entre todas las parejas de vértices reales de que se compone cada supervértice. El algoritmo asignará los arcos obtenidos a las parejas de vértices a menor distancia, pudiendo considerarse como función distancia una función puramente geográfica o una función combinada en la que intervengan también parámetros relativos al tráfico, grado de los vértices, etc. Siguiendo con el ejemplo anterior, en la Fig 7 queda representada la situación final tras la expansión de arcos.

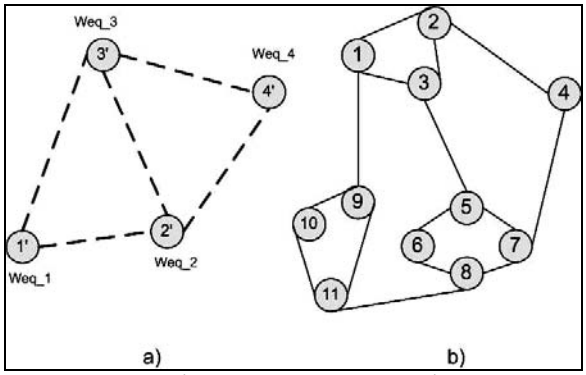


Figura 7: Aplicación del algoritmo de expansión a un ejemplo

3 Aplicación en la herramienta de diseño TAROCA-IP

Taroca-IP (Traffic Routing and Cost Análisis-IP) es un herramienta software diseñada e implementada por el Grupo de Ingeniería de Telemática (GIT) de la Universidad de Cantabria para la planificación, diseño y dimensionado de la capa lógica de las redes Internet de Nueva Generación, ver [12].

El algoritmo presentado en la sección 2 se ha incorporado en TAROCA-IP, integrándolo dentro de la etapa encargada de la optimización de la Estructura de Red, ver Fig. 8. Esta etapa de planificación está compuesta en TAROCA-IP a su vez por tres módulos: Clasificación, Distribución y Cálculo de la Topología. El primero de los módulos, Clasificación, determina la clasificación jerárquica de los nodos como nodos de acceso, Edge, Core o de Interconexión. El segundo de los módulos, Distribución, se encarga de llevar a cabo una primera aproximación de la demanda de tráfico que será enrutada posteriormente a través de la red, considerando tres clases de tráfico (elástico, semielástico e inelástico). El algoritmo que se ha presentado proporciona el Cálculo de la Topología, correspondiente al tercer módulo de la Etapa de Estructura de Red.

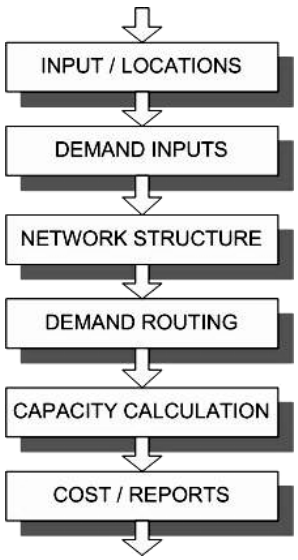


Figura 8: Descripción general en etapas de TAROCA-IP

El cálculo de la topología se aplica únicamente a la estructura dorsal (core) de la red IP ya que se asume que cada nodo de acceso está conectado a su nodo edge asignado a través de una estructura de estrella o bien en alguno casos estrella doble por razones de seguridad. El módulo de topología parte de una red core totalmente mallada cuyos enlaces tienen asignados unos determinados valores de tráfico, proporcionados a través de módulos previos. El objetivo del módulo es conseguir un red core bi-conexa en la que el número de saltos del camino más corto sea menor o igual que un umbral determinado a priori. En la siguiente sub-sección se describe el procedimiento de cálculo de la topología a través del método explicado en la sección 2, de manera que en una red real los nodos y enlaces se corresponden con los vértices y arcos mencionados a lo largo de la descripción del algoritmo.

3.1 Aplicación del Módulo de Topología al diseño de una red NGI

Sobre la red dorsal totalmente mallada (estado inicial), se aplica en primer lugar el algoritmo de reducción que se encarga de eliminar todos los enlaces cuyo tráfico no supere un umbral dado y considerando en todo momento la limitación en número de saltos de la red, de manera que los caminos entre nodos no nunca superen el máximo

número de saltos establecido. En el algoritmo de reducción, los enlaces son recorridos según un orden establecido, resultante de aplicar como criterio de ordenación una función de peso que combina tanto el flujo de tráfico como la longitud física de los enlaces. a través del parámetro de entrada λ según la expresión $\omega_i(\lambda) = \lambda \cdot \frac{x_i}{x_{máx}} + (1-\lambda) \cdot \left(1 - \frac{l_i}{l_{máx}}\right)$.

De cualquier modo, como resultado de esta primera parte se obtiene una topología inicial conexa pero no necesariamente bi-conexa. En el caso de que la topología obtenida de la aplicación del algoritmo de reducción no cumpla la bi-conectividad de sus enlaces, el procedimiento continúa con la aplicación de los algoritmos de partición y expansión. El peso equivalente utilizado para cada subred se basa en un criterio múltiple que considera aspectos como el número de nodos, el flujo total de tráfico, diámetro y grado de la subred en cuestión. El proceso global que abarca el módulo queda ilustrado en la Fig. 9.

La pareja de subredes que serán unidas por un nuevo enlace se seleccionan en primera instancia en términos de grado mínimo con objeto de minimizar el número de enlaces que serán añadidos.

En el caso de encontrarse ante subredes de igual grado, entran en juego el flujo de tráfico y el número de saltos máximo entre las mismas, diámetro, de tal manera que las subredes origen y destino del nuevo enlace maximizan el valor de la expresión $W_{eq}(\psi) = \psi \cdot \frac{D_{S_S - S_D}}{D_{máx}} + (1-\psi) \cdot \left(1 - \frac{L_{S_S - S_D}}{L_{máx}}\right)$

donde, $D_{S_S - S_D}$ representa la demanda entre las subredes y $L_{S_S - S_D}$ la longitud media entre las mismas y los máximos de demanda y longitud se refieren a valores máximos entre todas las posibles parejas de subredes existentes. Observar, que cada vez que se restaura un nuevo enlace, se modifican la red de subredes y por tanto los valores que intervienen en esta expresión son recalculados. Así mismo, ψ ,parámetro de entrada a Taroca-IP, establece el peso relativo de cada factor (demanda y longitud) en la selección de subredes, de tal manera que según sus valores y de acuerdo a la expresión anterior, se favorecerá la selección de subredes de mínimo grado y máxima demanda entre ambas y/o la selección de subredes de mínimo grado y mínima longitud entre ambas .

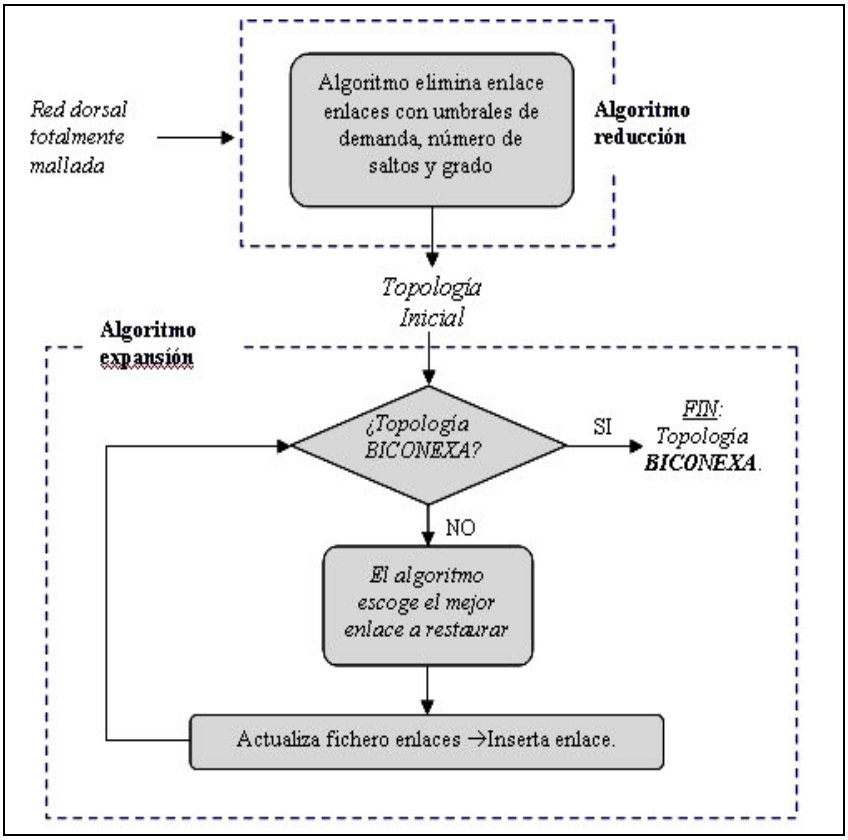
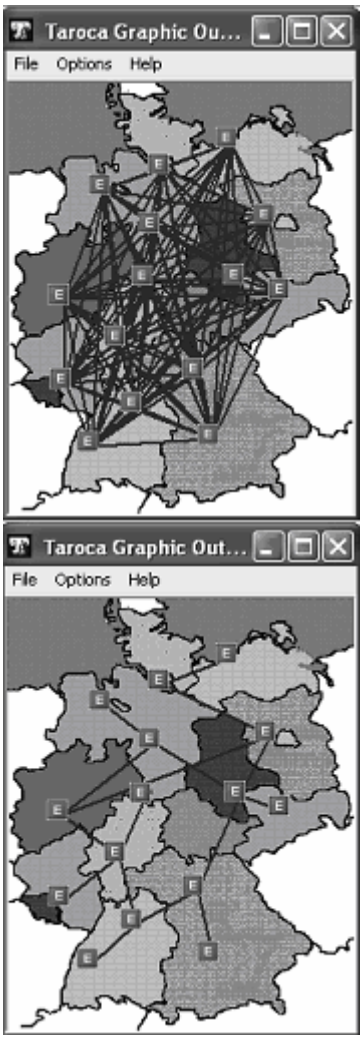


Figura 9: Descripción funcional del módulo de topología.

3.2 Aplicación en redes de prueba

Se ha estudiado el comportamiento práctico del procedimiento bajo un conjunto de ejemplos de redes reales basados en la aproximación de datos de redes core nacionales para el caso de España[4] y Alemania[5]. La solución ha sido evaluada en términos tanto de tiempo de ejecución como desde el punto de vista de la viabilidad práctica de una implementación real. Como última instancia, se ha utilizado también la representación gráfica de la solución obtenida para compararla con implementaciones reales de redes de proveedores de servicios de Internet.

En la Fig. 10, se muestra gráficamente el comportamiento del módulo de topología, donde partiendo de una topología totalmente mallada se obtiene una bi-conexa con limitación de 5 saltos.



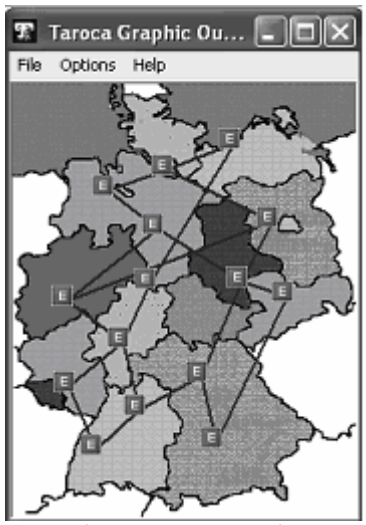


Figura 10: Topología para un caso hipotético de una red core IP de 15 nodos ubicada en el ámbito nacional de Alemania

Los resultados obtenidos en la Fig 10.b, tras la reducción de enlaces, muestran un topología con una subred principal bi-conexa en el núcleo y algunas subredes aisladas, algunas de ellas compuestas por un único nodo. Estos resultados se deben a la limitación de saltos y la fuerte demanda entre algunos de los nodos de la subred principal. La parte del módulo encargada de la expansión intenta conectar las subredes aisladas a través de los enlaces más cortos al nodo más cercano de la subred principal. De esta manera, se restauran cuatro enlaces y la topología obtenida asegura la bi-conectividad entre todas las parejas de nodos.

En la Tabla I se muestran los resultados obtenidos tras un estudio del comportamiento del tiempo de cómputo, realizado a través de la generación de redes completamente malladas con las ciudades más importantes de Alemania. Así mismo, en la Fig. 11 se representa el tiempo de cómputo con el número de nodos dorsales.

Tabla I: Tiempo de ejecución de los algoritmos [s]

Número de nodos	Tiempo Ejec. Reductivo	Tiempo Ejec. Expansivo
25	0.4	0.3
50	0.4	0.3
75	2.5	0.6
100	7	2
200	19	4.2
300	52	5.7
400	288	10.4

Como puede observarse, la mayor parte del tiempo de procesamiento se consume en la parte de reducción.. Esto se debe al hecho de que el número de enlaces a considerar se incrementa con N^2, donde N se corresponde con el número de nodos. Además, el control del número de saltos en esta misma parte requiere un tiempo proporcional a M*N, lo que se traduce en una dependencia de N^3 del tiempo con el número de nodos en esta primera parte del módulo.

Con todo, se concluye que el algoritmo es aplicable para el diseño de redes de medio a gran tamaño.

[4] Los datos del ejemplo de España se generan para los municipios más importantes y su demanda se calcula en base a su número de habitantes.

[5] El calculo del ejemplo de Alemania se basa en un estudio para la ANRT RegTP, Bonn , véase www.regtp.de.

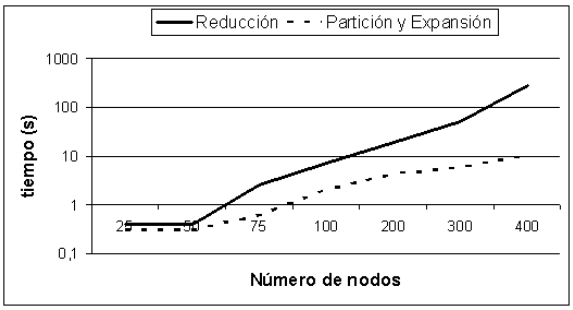


Figura 11: Tiempo de ejecución del algoritmo

4 Conclusiones y trabajo futuro

Este documento presenta un algoritmo para el cálculo de una topología bi-conexa considerando una restricción en el valor máximo de número de saltos de los camino cortos entre sus nodos. Se indica la aplicación del mismo con su incorporación en una herramienta de planificación para el diseño de redes core NGI.

Con los resultados obtenidos a través de los primeros estudios se deduce que este algoritmo proporciona resultados aceptables en estudios sobre redes reales de ámbito nacional o bien en redes IP europeas. El siguiente paso es estudiar el comportamiento del algoritmo bajo diferentes funciones de peso así como diferentes tipos de secuencias de ordenación en las partes correspondientes a la eliminación e incorporación de enlaces.

Este estudio va a ser realizado tanto en redes reales como en redes generadas de manera aleatoria, utilizando en su análisis diferentes tipos de criterios de evaluación como son la longitud total de la topología resultante, el número total de enlaces, grado medio de los nodos, tráfico de tránsito etc. y minimizando la función de coste $C_{total} = \sum_{e_i con\ x_i>0} c_1 \cdot x_i + c_2 \cdot l_i + c_3 \cdot x_i \cdot l_i$, con x_i flujo de trafico en el arco e_i y l_i su longitud. Adicionalmente, basándose en C_{total} está previsto comparar las soluciones obtenidas con el algoritmo con soluciones calculadas a través de la aplicación de modelos clásicos matemáticos y programas estándar para optimización mixta-entera, ver [13].

Agradecimientos

Este trabajo está avalado por la Red de Excelencia EURO-NGI del VI Programa Marco Europeo, IST-50/7613 y por el proyecto nacional TIC-2003-05061 promovido por el Ministerio Español de Ciencia y Tecnología.

Referencias

[1] M. Gondran and M. Minoux, “Graphs and Algorithms”, John Wiley and Sons, 1995.

[2] I. Ljubic and J. Kratica, “A genetic algorithm for bi-connectivity augmentation problem”, In C. Fonseca et al., editors, Proceedings of the 2000 IEEE Congress on Evolutionary Computation, pages 89-96. IEEE Press, 2000.

[3] Greg N. Frederickson, ”Ambivalent data structures for dynamic 2-edge-connectivity and k smallest spanning trees”, Proceedings of the 32nd annual symposium on Foundations of computer science, pages: 632 - 641. IEEE Computer Society Press, 1991.

[4] A. Zhu, ”A uniform framework for approximating weighted connectivity problems”, B.Sc. thesis, University of Maryland, MD, May 1999.

[5] Monika H. Rauch, ”Fully Dynamic bi connectivity in graphs”, Algorithmica 13(6), pages 503-538, Jun 1995.

[6] D. Jungnickel, ”Graphs, Networks and Algorithms”, Springer Berlin, 1999.

[7] H. Whitney, “Congruent graphs and the connectivity of graphs”, Amer. J. Math. 54, pages 150-168, 1932.

[8] G. N. Frederickson, “Data Structures for On-Line Updating of Minimum Spanning Trees”, SIAM Journal Computation 14, pages 781-798, 1985.

[9] S. Khuller, R. Thurimella, “Approximation algorithms for Graph Augmentation”, Journal of Algorithms 14(2), pages 214-225, 1993.

[10] R. E. Tarjan, “Depth First Search and Linear Graph Algorithms”, SIAM J. Comput. 1, pages 146-160, 1972.

[11] R. E. Tarjan, J. Westbrook, ”Maintaining bridge connected and bi-connected components on-line”, Algorithmica 7, pages 433- 464, 1992.

[12] K. D. Hackbarth et al., 1st Deliverable of the WP 3.4, “Development of an European Network Design Tool for Next Generation Internet”, www.euro-ngi.org.

[13] M. Pioro et al., 1st Deliverable of the WP 3.1, “Optimisation of protected multi-layer core networks: topology, layout, flow and capacity design”, www.euro-ngi.org .

Evaluation of Packet Scheduling Policies with Application to Real-Time Traffic

Agustín Santos, Antonio Fernández, Luis López,
{asantos,anto,llopez}@gsyc.escet.urjc.es
Laboratorio de Algoritmia Distribuida y Redes (LADyR)
Universidad Rey Juan Carlos
C/ Tulipán S/N
28933 Móstoles (Madrid)

Abstract *Well-known theoretical results have proven that, in worst-case situations, buffer scheduling disciplines based on the traditional FIFO policy (and other simple policies) may produce unstable networks even at very low network loads. When this happens, the total number of packets, and hence the latency experienced, tends to infinity as time grows. However, similar analyses have demonstrated the existence of novel disciplines that are stable (bounded queues and delays) independently of the topology, as long as the networks is not fully loaded.*

In this paper, we present a new methodology and a set of simulations whose objective is to compare the performance of all these policies in a realistic network scenario. Based on the obtained results, we conclude that stable novel disciplines seem to present better characteristics than FIFO and could have an important impact on the quality of service of multimedia and other types of real-time-constrained traffic.

1 Introduction

In packet networks, it is common to have several packets attempting to cross the same link at a given time instant. The criterion used to choose the packet that does so first is called the link's *packet scheduling policy or strategy.* It is well-known that the scheduling policies used at the routers of a communication network drastically influence the performance of the network. In order to study this influence in worst-case scenarios, there are two classical theoretical network models of adversarial injection of packets.

The first adversarial model was proposed by Cruz [4, 5], and assumes that all the packets in the network are grouped in sessions, with associated paths and arrival rates. It also assumes that the arrival (injection) of packets into the system is decided by an adversary. In this model it has been shown the existence of policies that guarantee stability [9, 10] as long as no link is fully loaded, while FIFO can be unstable [2].

In a second model, called Adversarial Queueing Theory (AQT) [3, 1], the adversary still controls the injection and paths of the packets, but these need not belong to any session. However, the system is assumed to evolved in discrete steps, which may imply that all packets have the same length and all links the same bandwidth. In this model, it has been shown the existence of policies that are stable for any underloaded network, and policies (FIFO among them) that can be unstable at *any constant load.*

Recently, a third model that tries to merge the above two has been proposed [7, 6]. It has been called Continuous AQT (CAQT), since it frees the AQT model of its synchronous behavior. In this model packets need not be grouped into sessions and may have different lengths, while each link has associated a bandwidth and a propagation delay. It has been shown in [7] that some of the stable policies in AQT are so in CAQT as well.

These theoretical results give a foundation that helps to face the problem of choosing an appropriate scheduling policy from a new perspective. Clearly, if we can make a choice, we should prefer a stable policy to a potentially unstable one. However, for latency-critical and real-time applications it is not enough to have guaranteed delays, but we need to have also small delays and low jitter. Unfortunately, the delay bounds that have been proven for stable policies are too large to be useful in real systems and there is no known bound for the jitter. Hence, we would like to know more about the real behavior of different scheduling policies in realistic set-ups.

1.1 Contributions

In this paper we start the evaluation of the simple scheduling policies considered for the AQT and CAQT models [3, 1, 7] in real environments by means of simulation. The final objective is to compare the results obtained with this empirical evaluation with those obtained analytically.

As mentioned, to start the evaluation we use simulations. We fix the network topology to an 11×11 torus, in which every node is, at the same

time, router, source, and sink of packets. We make each node to generate a new packet periodically, whose destination is chosen random and uniformly among the nodes of the network. The routing is deterministic, so that the traffic is balanced among all the links. We then vary the average load of the network in different simulations. To do so we vary the packet generation frequency. We are very interested on the response of the network with high levels of load.

As we said, in this set up we study several simple scheduling policies that have been studied analytically for the CAQT model [7]. These are, besides FIFO, the following. The Longest-in-System (LIS) policy gives higher priority to older packets, while the Shortest-in-System (SIS) policy gives higher priority to newer packets. The Farthest-from-Source (FFS) policy gives higher priority to the packets that have traversed larger number of links, while the Nearest-from-Source (NFS) policy gives higher priority to the packets that are closer to their source node.

In the simulation results obtained, especially at high loads, it can be observed, that from the above policies, the average delay suffered by a packet is clearly smaller with the FIFO, LIS, and FFS policies than with SIS and NFS. This initial result is surprising, since it is known that both SIS and NFS are stable in any network that is not fully loaded in the CAQT model, while FIFO and FFS are unstable in the same model for any constant load (i.e., for any constant load an appropriate network can be built in which these policies are unstable).

Then, when we look at the variance of these three policies we observe that the variance with FIFO and LIS is much smaller than that of FFS. This means that the former policies present smaller jitter. Furthermore, we observe that the variance and the maximum observed delay (i.e., the jitter) are significantly smaller with the LIS policy than with FIFO. These results, and the fact that LIS is universally stable, imply that LIS should be considered in the future as an alternative to FIFO to be used by routers for real-time traffic.

The rest of the paper is organized as follows. In Section 2 we present the model assumed in the evaluation performed. In Section 3 we briefly review the simulation program we have used. In Section 4 we describe the specific simulations we have done and present some of the results obtained. Section 5 is devoted to a discuss and analyze these results in order to identify the policies that are best suited to real-time and multimedia traffic. Finally, in Section 6 we present conclusions and future lines of work.

2 Model description

Our model is strongly inspired by Cruz's, the AQT, and the CAQT models. In these models a network is modelled as a directed graph in which the nodes are simultaneously transmitting and receiving stations, and routers. The edges of the graph are the (unidirectional) links of the network. In CAQT, which is the most general of these models, the communication system is controlled by a bounded adversary that chooses routes and injection times for all the packets. This adversary is restricted by a load parameter $r \leq 1$ and a burst parameter $b \geq 1$. Then, for any link e in the network with bandwidth B_e, and any interval of time I, the length of all the packets injected in the interval I that need to cross e cannot add up to more than $r|I|B_e + b$ bits. This is a classical "token bucket" restriction [13].

As we mentioned above, in our simulations, we assume the network topology to be a torus with $N \times N$ nodes (in particular, we use an 11×11 torus). In this topology, each node is connected to four nodes, which we call north, south, east, and west. Hence, nodes are connected by directed edges to their neighbors in their same row and column. Note that we have two edges connecting adjacent nodes, one for each direction. Hence, the result is an isotropic fully symmetrical topology based on degree-4 point to point interconnections. We have chosen this kind of topology because it is rich enough to present diverse path lengths and routes and, at the same time, it remains simple to model and analyze. This kind of toroidal mesh has been deeply studied in the literature and its simple rectangular structure and scalability are well-known. It has been used as a means for connecting computation elements in parallel computers [8].

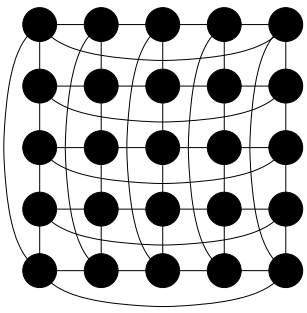


Figure 1: A 5×5 torus. Each picture edge represents two unidirectional links with opposite directions. Observe that this structure is completely isotropic and symmetric, in the sense that all nodes are topologically equivalent.

We make standard assumptions about the network for our analysis. In our model, all the packets will have the same length. Packets are generated at each node at a periodical rate (in packets per second) which will determine the load of the network. The destinations of the packets

are uniformly distributed over all the nodes in the network. (This is one of the differences between our model and the above mentioned model, in which the adversary chooses the destination.) All links have the same capacity (bandwidth) in bits per second. Only one packet can be transmitted across an edge at any time, although nodes can receive and transmit more than one packet at any instant. Packets are buffered when the next edge in their path is busy. In each router, the packet buffer for each outgoing link can hold an unlimited number of packets. As we said, we evaluate scenarios in which at the link buffers the packets are scheduled according to the following queueing policies: First-In-First-Out (FIFO), Longest-In-System (LIS), Shortest-In-System (SIS), Nearest-from-Source (NFS), and Farthest-from-Source (FFS). For each scenario, all routers use the same queueing policy.

The routing of packets in our network is deterministic. There is a unique path to reach a given destination from a given node. Furthermore, this path is guaranteed to be a shortest path. (Note that the routing used again makes our simulation scenarios different from the model considered in the above mentioned models, in which the adversary gets to choose the packet, or session, paths and can make packets to traverse paths that are not shortest in the network.) Additionally, this set of paths is chosen so that each link in the network is contained in a subset of the same size and types, and hence the traffic is balanced among all links. A greedy routing algorithm has been developed for these simulations and inserted in the simulator. The algorithm can be used for any $N \times N$ torus. The following is a brief description of the routing algorithm implemented. At each router, each destination has a preferred outgoing link for forwarding packets addressed to that destination. The preferred outgoing links for all the destinations are distributed evenly among the four outgoing links in each router. This is done by dividing the nodes in four quadrants, centered at the four axes determined by the links, and assigning each quadrant to its corresponding link. Then, at each node, a packet is always sent along the preferred outgoing link for its destination at that node.

3 Simulation environment

All the simulations in this article have been carried out using J-Sim [11] (formerly known as JavaSim). J-Sim is a component-based, compositional simulation environment which offers a wide variety of dispatching rules and in which new rules can be easily specified. It has been built upon the notion of the autonomous component programming model. The program runs on virtually any computer supporting Java. J-Sim has been designed to simulate network behaviors in a realistic way. This means that it tries to take into consideration any kind of variable which could have an impact in real scenarios, including propagation delays, packet processing times, etc. Like ns-2 [12], J-Sim is a dual-language simulation environment in which classes are written in Java (for ns-2, in C++) and glued together using Tcl/Java. However, unlike ns-2, classes/methods/fields in Java need not be explicitly exported in order to be accessed in the Tcl environment. Instead, all the public classes/methods/fields in Java can be accessed (naturally) in the Tcl environment.

We have modified the J-Sim package in several ways, mainly to adapt it to our model. First, the traffic generator has been modified in order to ensure that destinations are uniformly distributed over all nodes in the network. Then, the sink monitor has also been changed in order to log several parameters that are not stored by J-Sim (e.g., the mean and the variance of the packet delay and the queue size, and samples of these values chosen at random). As we have mentioned, the routing protocol has been replaced in order to implement the routing algorithm discussed above.

4 Simulations

We have run a number of simulations, in different simulation scenarios. Each simulation scenario is defined by a scheduling policy and a traffic load. In total, we have studied 25 scenarios corresponding to the five queueing policies we introduced previously and to five different traffic loads. For each scenario several simulations have been run to verify that the measures obtained are basically the same.

The traffic load has been controlled by changing the periodicity (rate) of packet generation at the nodes. We first compute the rate λ (in packets per second) at which nodes need to generate packets in order to reach a 100% expected utilization of all the communication links of the network. Then, we scale down this rate to force the desired average load. To derive the rate λ we first observe that in the $N \times N$ torus, for N odd, given a node, there are $4i$ nodes at distance $i \in \{0, ..., \frac{N-1}{2}\}$, and $4(N-j)$ nodes at distance $j \in \{\frac{N+1}{2}, ..., N-1\}$. Then, we can easily compute the expected length of a packet path, $E[PL]$, since the destination of a given packet is chosen uniformly at random, as

$$\begin{aligned} E[PL] &= \frac{1}{N^2}\left(\sum_{i=0}^{\frac{N-1}{2}} 4i^2 + \sum_{j=\frac{N+1}{2}}^{N-1} 4(N-j)j\right) \\ &= \frac{N^2-1}{2N}. \end{aligned}$$

There are N^2 nodes generating packets, and since there are $4N^2$ unidirectional links and the load is balanced among them, we obtain that the expected load of an edge (in packets per second)

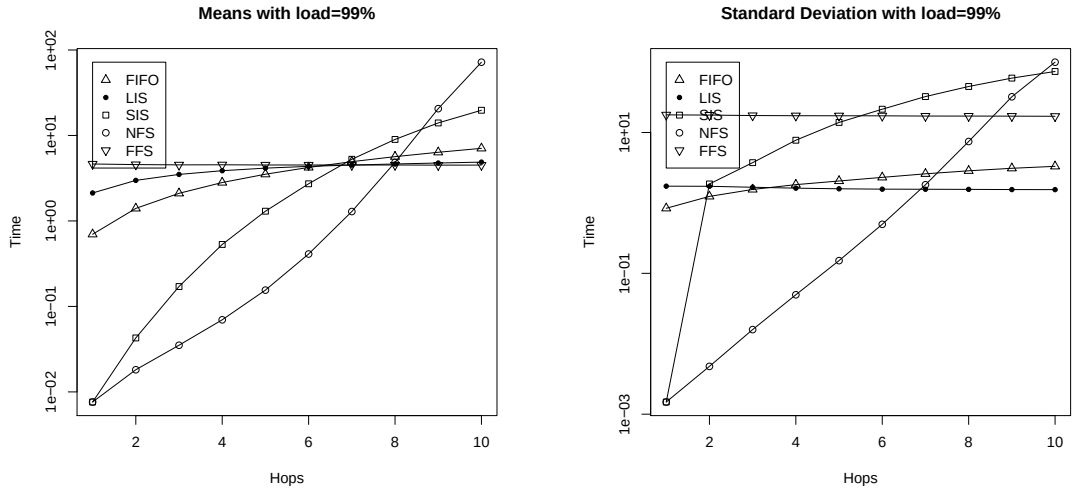


Figure 2: Average and standard deviation of the latency experienced by the packets for a network load of 99%. As it can be observed, the scheduling policy being used has a deep impact on the delay and jitter of packets.

for a generation rate λ is $\lambda N^2 \frac{N^2-1}{2N}/(4N^2) = \lambda\frac{N^2-1}{8N}$. We assume that all packets have a length of L bits and all links have a bandwidth of B bits/second. Since a link can process B/L packets per second, a link is 100% loaded for

$$\lambda = \frac{8NB}{(N^2-1)L}.$$

Then, if we want a load in the network of $0 < r < 1$, we will make each node to generate packets at a rate of $r\lambda$ packets per second.

As we mentioned, we are interested on the behavior of the system for high loads. For that reason, we have used the following five different values of network load in our simulations: 99%, 97%, 95%, 90%, and 50% (i.e., $r \in \{0.99, 0.97, 0.95, 0.90, 0.5\}$). The packet size has been fixed to 512 bytes, whose length becomes 532 bytes with the IP protocol headers (i.e., $L = 532 \times 8$ bits). The bandwidth of all the links has been fixed to $B = 835200$ bits per second. Once a particular scenario (pair policy-load) has been chosen, we perform a simulation during $6,000$ seconds on the 11×11 torus. The simulation results are dumped in a log file in which the information of the packets is classified depending on the number of links it has traversed to reach its destination (number of hops within the network). For example, a hop count of one for a given packet means that the packet traversed only one link (the destination was a neighbor of the source).

5 Simulation results

Average delay and delay jitter are two of the most important Quality-of-Service (QoS) parameters when dealing with the transmission of real-time multimedia traffic. It is well known that the average delay is related to the mean time packets spend in travelling from source to destination, while the delay jitter measures the dispersion on the value of the delays of different packets belonging to the same session. The delay jitter can be measured in several ways, two of the most appropriate ones are through the standard deviation of the packet delay distribution or through the maximum difference between any two given delays of the session.

These parameters have been extensively studied in the preceding literature [13] and its impact on real-time traffic has been clearly established both in ATM technologies, where some guarantees may be supplied, and in best-effort IP networks like the Internet, where important efforts have been carried out trying to offer improved QoS for the support of multimedia protocols.

It is clear that average delay is specially important in applications involving some kind of user interaction like VoIP, whereas jitter is the dominant factor in non-interactive video or audio streaming over the Internet. For these reasons, we concentrate on measuring the behavior of these two parameters on the different simulation scenarios we have presented.

Out of the results obtained, we show in Figs. 2, 3, and 4 the measured values of the average and standard deviation of the latencies experienced by the packets for network loads of 99%, 90%, and 50%, respectively, with the five policies we introduced previously. These curves show the values classified by the different number of hops traversed by the packets. The lower the number of hops, the shorter the trip of packets on the network and vice-versa. In these figures, it is interesting to observe that FFS, LIS, and FIFO are the policies

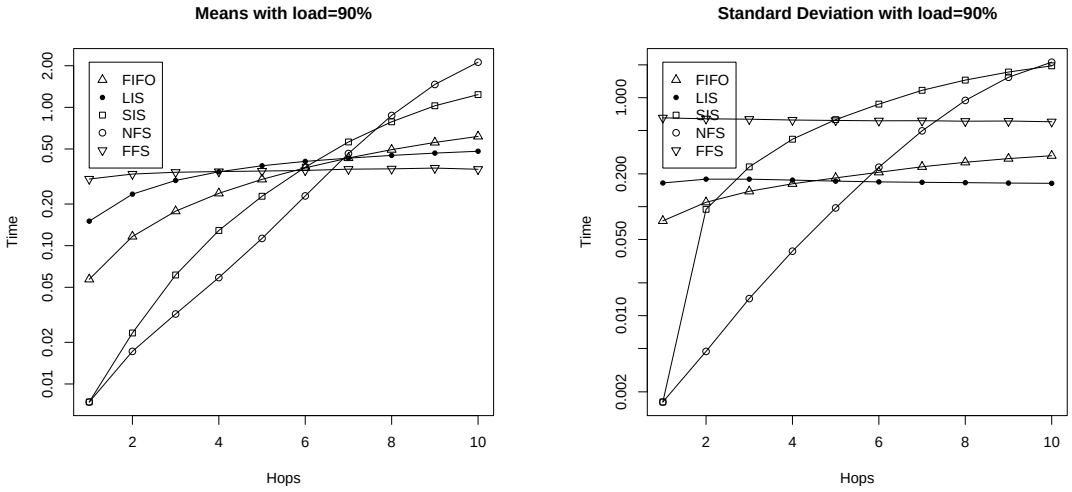


Figure 3: Average and standard deviation of the latency experienced by the packets for a network load of 90%.

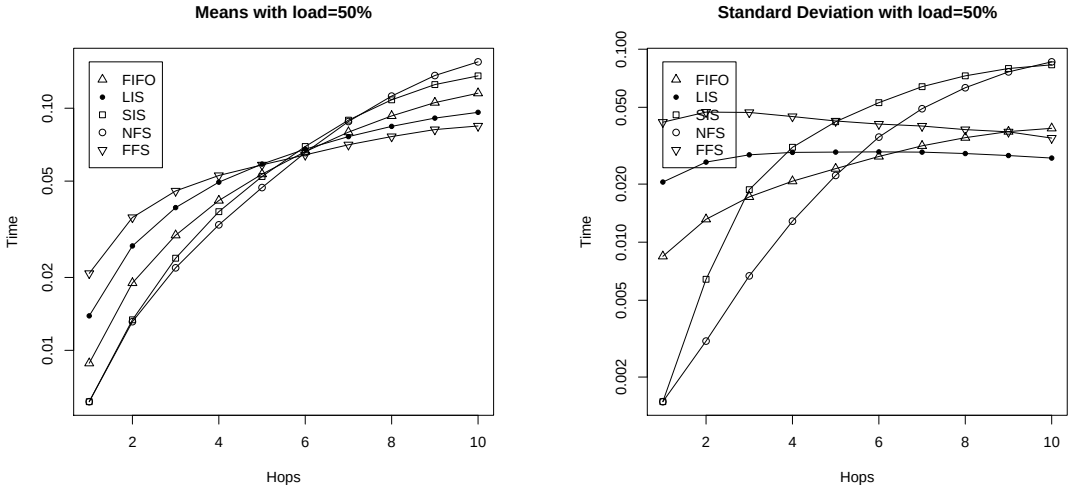


Figure 4: Average and standard deviation of the latency experienced by the packets for a network load of 50%.

with the best mean latency for the higher number of hops, especially for high network load, and these values remain rather constant as the number of hops varies. In fact, the other two policies, SIS and NFS, show an extremely large average latency for the packets that traverse 10 hops as the load increases. This result is somewhat surprising, since in the CAQT model SIS and NFS have been shown to be stable in any network as long as the link load is below 100% (i.e., $r < 1$), while FIFO and FFS have been shown to become unstable at any constant load (and for FFS, this is done in a torus-like network [3]).

When we look at the standard deviation, we can observe, again, that SIS and NFS show the largest values for high number of hops. Among the other three policies, we see that the deviation of FFS is also significantly higher than that of FIFO and LIS. Since the standard deviation of the latency has a lot to do with the jitter experienced, we discard FFS from further consideration.

Then, we center our attention to compare FIFO with LIS. Fig. 5 presents the average latency and the standard deviation experienced by packets for the different network loads under these two policies, independently of the number of hops traversed. This figure shows that, while FIFO has a lower average latency, it also has a larger standard deviation, and hence a larger jitter level. This can be observed more clearly in Fig. 6, where LIS shows a maximum latency that is extremely stable and significantly lower than that of FIFO for high number of hops.

A second interesting observation from Fig. 6 is

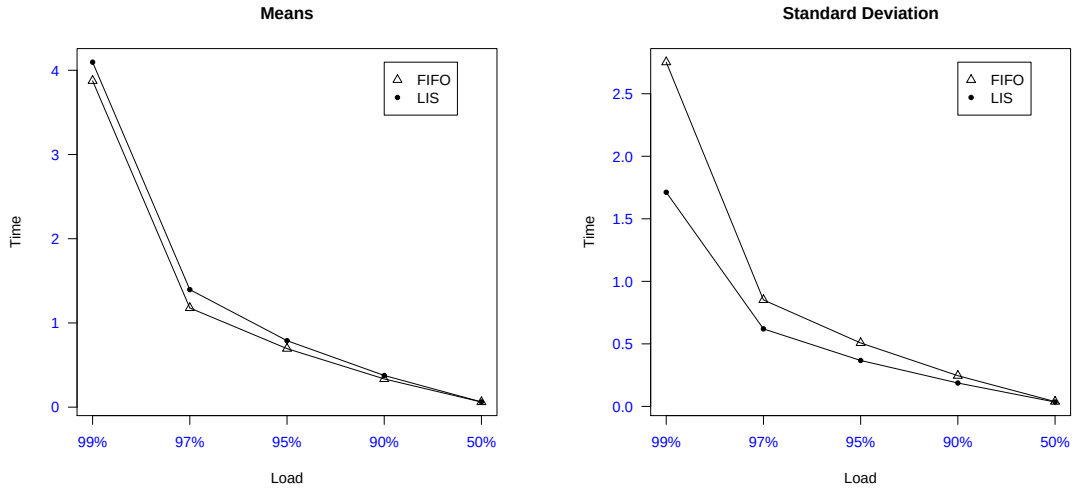


Figure 5: Average and standard deviation of the latency experienced by the packets with FIFO and LIS for different network loads. The depicted values have been calculated considering all possible network distances.

that FIFO has a larger average latency than LIS for high number of hops. Furthermore, the average latency with FIFO seems to follow a linear pattern, growing steadily with the number of hops, while the same metric with LIS seems to converge to a value around 5 seconds. A similar fact can be observed also in Fig. 7, where it can be seen that the latency experienced by packets is much more concentrated (around 5 seconds) for LIS than for FIFO.

Observing these results, we can claim that, at least for the considered network topology, the policy based on the LIS algorithm is more suitable than FIFO for obtaining an improved QoS on streaming applications requiring low jitter. Moreover, the high stability presented by LIS curves, make it more appropriate for that kind of applications involving real time interaction between multiple users placed at different network locations. These applications include video-conferencing, voice-chat-rooms, group IP telephony, network games, etc. The justification for this can be easily understood observing Fig. 8. Using FIFO, both the latency and the jitter strongly depends on the distance (in terms of hops) from source to destination. Hence, the quality of the session depends highly on the proximity to the source. On the other hand, in LIS, this difference is minimized, and both parameters remain stable with distance, guaranteeing an equivalent quality on all receivers.

6 Conclusion

In this paper, we have imported the main ideas and results from the CAQT theoretical model to develop a set of simulations comparing the performance of the FIFO queue scheduling discipline with other policies, which have been proved to present better stability properties in the framework of this model. With this objective, our research has been focused on designing this simulations for a $N \times N$ toroidal mesh, where we establish a uniform set of random sessions and measure the two main parameters involved in the QoS: the delay and jitter of packets. The obtained results suggest that, at least for that topology, the LIS policy can provide significant advantage, with respect to FIFO and other policies, for real-time-traffic transmission.

This result opens a whole new research domain where further investigation should be carried out to generalize the presented conclusions and to design novel mechanism suitable for implementing this policy in a real environment. In particular, new packet generation algorithms (random uniform, exponential, etc.) and more queueing policies should be tested. Besides, larger and more complex network topologies should be explored (scale free, hierarchical, etc.), including those where different scheduling mechanism coexist. Finally, novel prototypes should be developed to verify experimentally the results obtained though simulation.

Acknowledgements

This work has been partially supported by URJC under grant PPR-2004-42.

References

[1] M. Andrews, B. Awerbuch, A. Fernández, J. Kleinberg, T. Leighton, and Z. Liu. Uni-

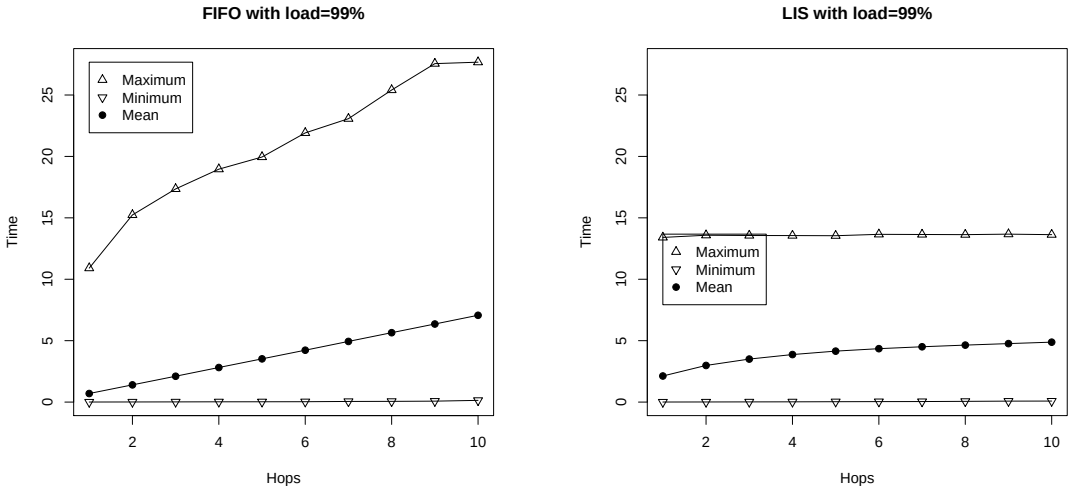


Figure 6: Average, minimum and maximum latency experienced by packets with FIFO and LIS for 99% network load.

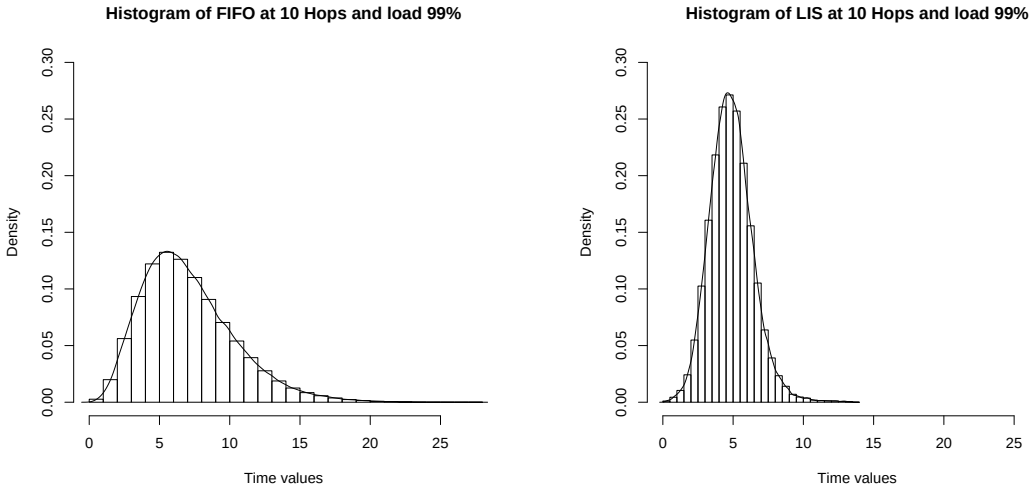


Figure 7: Histogram with the density distribution of latencies for FIFO and LIS with 99% network load. As it can be observed, the dispersion on the FIFO delay is clearer over the LIS one.

versal stability results for greedy contention–resolution protocols. *Journal of the ACM*, 48(1):39–69, 2001.

[2] Matthew Andrews. Instability of fifo in session-oriented networks. *J. Algorithms*, 50(2):232–245, 2004.

[3] A. Borodin, J. Kleinberg, P. Raghavan, M. Sudan, and D. Williamson. Adversarial queueing theory. *Journal of the ACM*, 48(1):13–38, 2001.

[4] R.L. Cruz. A calculus for network delay. Part I: Network elements in isolation. In *IEEE Transactions on Information Theory*, volume 37, pages 114–131, 1991.

[5] R.L. Cruz. A calculus for network delay. Part II: Network analysis. In *IEEE Transactions on Information Theory*, volume 37, pages 132–141, 1991.

[6] J. Echagüe, V. Cholvi, and A. Fernández. Universal stability results for low rate adversaries in packet switched networks. *IEEE Communication Letters*, 7(12):578–580, 2003.

[7] María J.Blesa, Daniel Calzada, Antonio Fernández, Luis López, Andrés L. Martínez, Agustín Santos, and María J. Serna. Adversarial queueing model for continuous network dynamics. Submitted, 2005.

[8] F. T. Leighton. *Introduction to Parallel Algorithms and Architectures: Arrays, Trees, and*

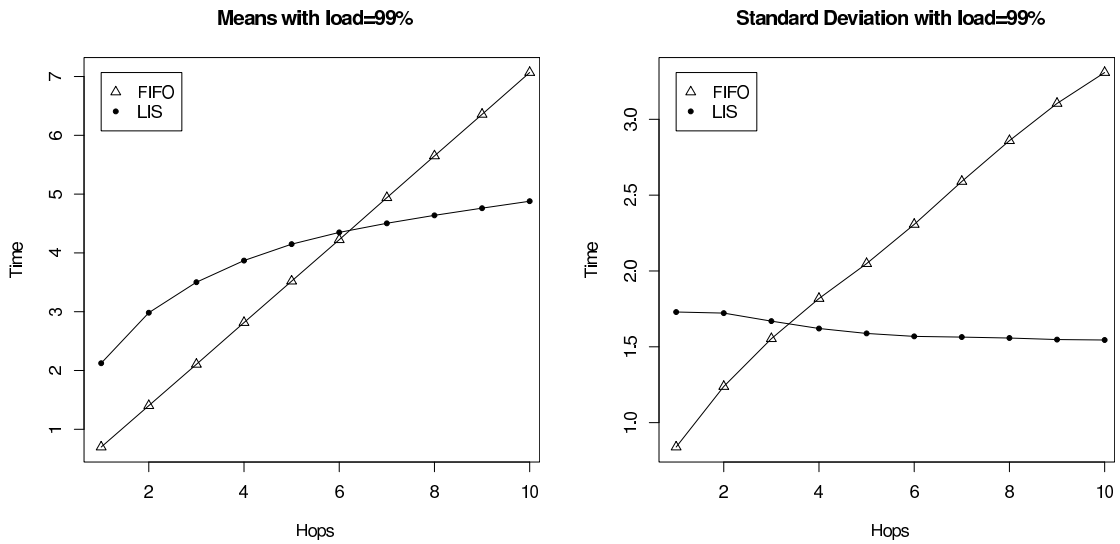


Figure 8: Average delay and standard deviation for FIFO and LIS policies as a function of the distance (in number of hops) from source to destination. The simulations have considered a network load of 99%.

Hypercubes. Morgan Kaufmann, San Mateo, 1992.

[9] A. K. Parekh and R. G. Gallager. A generalized processor sharing approach to flow control in integrated services networks: The single-node case. *IEEE/ACM Transactions on Networking*, 1(3):344 – 357, 1993.

[10] A. K. Parekh and R. G. Gallager. A generalized processor sharing approach to flow control in integrated services networks: The multiple-node case. *IEEE/ACM Transactions on Networking*, 2(2):137 – 150, 1994.

[11] J-Sim simulator. http://www.j-sim.org/.

[12] NS-2 simulator. http://www.isi.edu/nsnam/ns/.

[13] Andrew S. Tanenbaum. *Computer Networks*. Prentice Hall PTR, 4 edition, 2003.

Combinación de Mecanismos para la Mejora del Rendimiento de TCP sobre Canales Inalámbricos con Pérdidas a Ráfagas

Marta García, Ramón Agüero, Luis Muñoz, José Ángel Irastorza
Departamento de Ingeniería de Comunicaciones
Grupo de Ingeniería Telemática. Universidad de Cantabria
39005 Santander
E-mail: [marta, ramon, luis, angel]@tlmat.unican.es

***Abstract**. This paper shows an important enhancement of TCP behavior over the IEEE 802.11b Wireless Local Area Network (WLAN) if a combination of a Forward Error Correction (FEC) scheme and a Snoop agent is added to the idle Repeat reQuest (RQ) mechanism inherently used by such technology. The FEC scheme has been improved with some smart cross-layer optimizations, and its operation adapts to the channel conditions. Besides, the aim of the Snoop agent is to back up the FEC, overcoming the deficiencies of its adaptation, which are mainly due to the high variability of the channel inside a typical office environment. A complete experimental approach has been followed, and exhaustive measurement campaigns have been carried out over a real platform to derive a large number of performance parameters. Both methods have been implemented as modules belonging to a generic layer-two Performance Enhancing Proxy (PEP).*

1 Introducción

En los últimos años han proliferado los mecanismos destinados a la mejora del comportamiento de los protocolos de Internet sobre las redes de área local inalámbricas o tecnología WLAN. En particular, el rendimiento del protocolo de control de la transmisión (TCP, Transmisión Control Protocol) en el entorno de las redes sin hilos sigue siendo un campo de constante investigación. Tradicionalmente, las distintas implementaciones del protocolo TCP se han ido adaptando a las particularidades de las redes cableadas en las que las pérdidas de paquetes se deben principalmente a la congestión de los nodos intermedios. Sin embargo, las características intrínsecas de los canales inalámbricos pueden hacer ineficientes dichas implementaciones cuando se utilizan sobre tecnología WLAN, pudiéndose producir una degradación en el rendimiento debida, entre otros factores, a los tiempos de inactividad que se producen en la entidad TCP transmisora como consecuencia de los paquetes perdidos por culpa de las imperfecciones del medio radio.

En esta línea, surgen constantemente nuevas propuestas para mejorar el comportamiento de este protocolo en estas redes, propuestas que básicamente podemos dividir en dos grupos, dependiendo de si suponen una modificación del protocolo TCP en los equipos finales, o bien, si se solventan las pérdidas a nivel local, de forma transparente a los niveles superiores. La solución que se propone en este artículo corresponde al segundo grupo pero va más alla de otros trabajos [1], dado que combina dos técnicas de nivel de enlace como son la corrección de errores hacia delante (FEC) y las retransmisiones locales mediante un agente Snoop, conocedor del funcionamiento del protocolo TCP. Además de esto, la importancia del trabajo reside en su carácter exclusivamente experimental, con una caracterización exhaustiva del comportamiento de la propuesta, que ha sido implementada sobre una plataforma real en la que se utilizado la tecnología de WLAN más extendida hoy en día, como es el estándar IEEE 802.11b.

2 Técnicas de Mejora de Nivel de Enlace

En lugar de plantear una modificación en el protocolo TCP como se plantea en [2], este trabajo utiliza una propuesta local [3],[4], para hacer frente al impacto de los errores en el rendimiento de este protocolo. En particular, dicha propuesta consiste en añadir dos conocidas técnicas de nivel de enlace al mecanismo de retransmisión implícita definido en el estándar IEEE 802.11. Por tanto, este trabajo analiza la combinación de tres mecanismos diferenciados, que se describen brevemente a continuación.

2.1 Mecanismo de Retransmisión "Propietario"

El estándar IEEE 802.11 define un procedimiento de retransmisión implícita para hacer frente tanto a las colisiones como a los errores producidos por el canal. Cuando una estación recibe una trama, realiza el correspondiente chequeo de redundancia cíclica (CRC, Cyclic Redundancy Check). Si la trama recibida es correcta enviará un reconocimiento a la estación transmisora; en caso contrario, el transmisor procederá a la retransmisión de la trama una vez expirado un cierto temporizador. El término "propietario" que se ha añadido a esta mecanismo pretende constatar el hecho de que la mayoría de las implementaciones comerciales o adaptadores de red

que existen en el mercado no permiten controlar las retransmisiones 802.11 de manera que este mecanismo es completamente cerrado e inaccesible [5]. Este es el caso de los adaptadores 802.11b de Orinoco empleados en este trabajo, que utilizan tres reintentos de transmisión por trama, no siendo posible acceder a este parámetro ni modificarle. Este hecho, como se explicará en detalle más adelante, tiene sus implicaciones a la hora de llevar a cabo la implementación de la solución propuesta. Por tanto, la recepción de cuatro tramas consecutivas erróneas supone la pérdida de un paquete TCP, es decir, la tasa de error de trama (FER, Frame Error Rate) no es lo mismo que la tasa de pérdida de paquetes, que se puede definir, igualmente, como pérdida MAC (Medium Access Control) residual.

2.2 FEC Adaptativo

El principal objetivo de un esquema FEC a nivel de enlace es la corrección del mayor número posible de las tramas erróneas que llegan a un receptor. Su diseño necesita una exhaustiva campaña de medidas previa, con la que caracterizar la distribución de los bits erróneos por trama. Tras llevar a cabo dicha caracterización sobre un canal con una baja relación señal ruido (SNR), canal que será descrito posteriormente, se elige un código Reed Solomon acortado (750,720) definido sobre un cuerpo de Galois GF (2^{16}) con una capacidad de corrección de 15 símbolos. Teniendo en cuenta el número de bits erróneos por trama y el tamaño medio de la ráfaga de bits incorrectos, este código debería corregir alrededor del 70% de las tramas erróneas para una MTU (Maximum Transfer Unit) de 1500 bytes. Este esquema FEC introduce una sobrecarga de 480 bits por trama y se aplica sobre el datagrama IP completo.

Con objeto de minimizar el coste computacional introducido por este esquema se han seguido dos estrategias. Por un lado, el proceso de decodificación hace uso de la capacidad detectora del CRC del estándar IEEE 802.11. Si una trama se recibe sin error según el chequeo de CRC exportado por la tarjeta inalámbrica, el FEC extraerá la redundancia sin llamar a la función de decodificación. Por otro lado, para evitar que las tramas sean codificadas cuando las condiciones del canal son buenas, la información de SNR es exportada por la tarjeta inalámbrica y se utiliza para decidir si se protege la trama o no. Para elegir el umbral de codificación se ha caracterizado el canal buscando la distribución de tramas erróneas en función de la SNR observada. En particular, cuando la SNR baja por debajo de 15 dB, las tramas serán protegidas por el FEC; por el contrario, cuando la SNR mejore y supere los 18 dB, las tramas no llevarán ninguna redundancia.

Es necesario destacar que el mecanismo de retransmisión descrito en el apartado anterior y el esquema FEC operan de forma independiente, dado el carácter "cerrado" del primero de ellos. Si el FEC es capaz de recuperar una determinada trama, el transmisor no será consciente de ello y continuará retransmitiendo la trama hasta que reciba un reconocimiento porque ha llegado sin errores, o se alcance el límite máximo de retransmisiones. Por ese motivo, muchas retransmisiones serán inútilmente decodificadas, lo que provoca una pérdida de eficiencia. Para eliminar los datagramas duplicados que se generan, se utiliza el campo "identificador", presente en la cabecera IP.

2.3 Agente Snoop

El protocolo Snoop [6] de Berkeley es una técnica específica del protocolo TCP en la cual se introduce un agente Snoop en el punto de acceso (AP, Access Point) que actúa como nodo intermedio entre el terminal móvil (MT, Mobile Terminal) y la red cableada. Este agente monitoriza cada segmento TCP que atraviesa el AP en ambas direcciones y mantiene la información de estado de cada conexión. Asimismo, opera de forma diferente dependiendo si se trata del flujo de datos o de reconocimientos así como de tráfico de subida (desde el MT) o de bajada (hacia el MT).

Para el tráfico de bajada, es decir, desde un servidor de la parte cableada (FH, Fixed Host) hacia el MT, el agente Snoop guarda copias de los segmentos de datos que recibe del FH y los reenvía hacia la entidad TCP receptora en el MT. Por otro lado, el agente monitoriza los correspondientes reconocimientos enviados por el MT, borrando de la memoria los segmentos de datos que hayan sido confirmados. Además reacciona a la pérdida de paquetes en el canal radio realizando retransmisiones de los mismos que son disparadas, bien por la expiración de un temporizador local, o bien por la recepción de reconocimientos duplicados. Con objeto de prevenir que el FH invoque los procedimientos típicos de control de congestión de TCP, el agente también se encarga de suprimir los reconocimientos duplicados en el camino del MT hacia el FH.

Por otro lado, para el tráfico de subida, es decir, desde el MT hacia un FH, el agente Snoop detecta la pérdida de paquetes en el enlace inalámbrico monitorizando los segmentos de datos que él reenvía. Si dicha situación se produce, notifica al MT que la pérdida no ha sido causada por congestión, para que no dispare los correspondientes procedimientos para su control. En este caso se requieren modificaciones en la implementación de TCP en el MT como por ejemplo las opciones NAK [6] o ELN (Explicit Loss Notification [7]. Estas opciones no se han considerado en este trabajo dado el carácter local de la propuesta realizada que no contempla modificaciones en las implementaciones TCP.

3 Proxy Genérico para la Mejora del Rendimiento

El aspecto más innovador de la propuesta descrita en este trabajo se encuentra en el hecho de que las

técnicas mencionadas en la sección anterior operan de forma simultánea. Esto se consigue mediante un esquema basado en un proxy para la mejora del rendimiento (PEP, Performance Enhacing Proxy), situado entre la capa IP y la tecnología inalámbrica IEEE 802.11b, denominado Capa de Adaptación Inalámbrica (WAL, Wireless Adaptation Layer), [8]. La capa WAL trata de compensar el pobre rendimiento de los protocolos TCP/IP cuando operan sobre redes de área local inalámbricas con baja SNR, ocultando las deficiencias del canal radio a las capas superiores. Dicho proxy incorpora los módulos FEC y Snoop que funcionan como "protocol boosters". Un protocol booster es un elemento que mejora, de manera transparente, el rendimiento de protocolos superiores [3] de forma que el protocolo que es mejorado no varía su semántica extremo a extremo. La comunicación entre entidades WAL se basa en la definición de canales lógicos con un doble propósito: por una lado, el de aplicar los módulos acordes con el tipo de tráfico de que se trate y por otro lado, el de adaptar el comportamiento de los módulos al estado del canal medido en términos de SNR. Es evidente que dicho estado puede variar durante la comunicación entre las dos entidades. Mientras que el funcionamiento del agente Snoop no depende de las condiciones del enlace, en el caso del módulo FEC, puede resultar beneficioso adaptar su operación teniendo en cuenta el valor de la SNR observada. El cambio en la capacidad de corrección de la técnica FEC aplicada se lleva a cabo mediante el establecimiento de un nuevo canal lógico, tras el intercambio de algunos paquetes propietarios WAL que no suponen una gran sobrecarga.

Como módulo asimétrico, el agente Snoop sólo está activo en el AP, pero no en los MTs. Cuando el tráfico es de bajada (desde el AP hacia el MT), el módulo FEC es el último que visita el datagrama IP antes de ser entregado al controlador inalámbrico. Por el contrario, cuando el tráfico es de subida (desde el MT hacia el AP), el paquete pasa primero por el módulo FEC, por lo que el agente Snoop sólo recibirá datagramas libres de error.

4 Plataforma de Medidas y Definición de Parámetros

Para validar la arquitectura propuesta, se llevó a cabo una completa campaña de medidas en una plataforma experimental. Esta sección resume los aspectos principales de la misma. Adicionalmente se definirán los parámetros que se emplearán más adelante para verificar las mejoras de la propuesta.

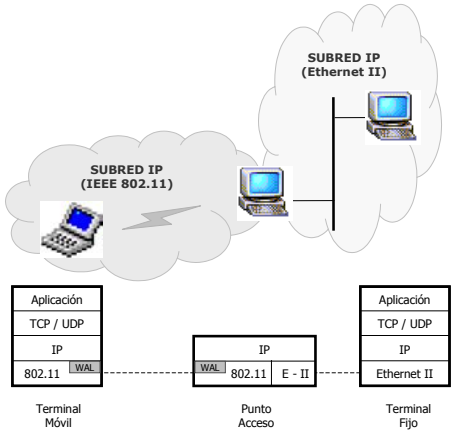


Figura 1. Plataforma de medidas experimental

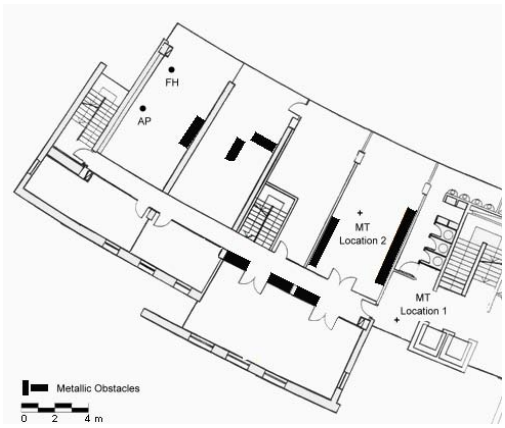


Figura 2. Entorno empleado durante la campaña de medidas

La plataforma experimental, que se muestra en la Figura 1, consta de tres equipos Pentium III, con sistema operativo Red Hat Linux 7.1 (versión de kernel 2.4.9). El Terminal Móvil (MT), actuando como cliente FTP, descarga un fichero de 10 Mbytes desde un servidor de ficheros o terminal fijo (FH), atravesando el punto de acceso (AP). El resto del tráfico de la parte cableada se considera despreciable. Tanto el AP como el MT incorporan tarjetas PCMCIA IEEE 802.11b de Orinoco configuradas en modo ad hoc. El cliente se situará en dos posiciones diferentes, dentro de un entorno típicamente de oficinas, tal y como se puede ver en la Figura 2. Para adaptar la operación del FEC a la calidad del enlace radio, el AP manda un mensaje propietario (denominado "beacon") cada 5 segundos, al que el MT responderá, reportando la SNR medida. Durante toda la campaña de medidas se utilizó la implementación de TCP Reno, con las opciones de "Timestamp" y reconocimiento selectivo (SACK, Selective Acknowledgement) activadas.

Para cuantificar la mejora introducida por el esquema propuesto se emplearán distintos parámetros que se describen brevemente a continuación:

Rendimiento o throughput: número de bits útiles recibidos (es decir, el tamaño del fichero en bits) dividido entre la duración de la transferencia.

FER: porcentaje de tramas MAC recibidas con error de CRC frente al total de tramas recibidas.

Longitud Media de Ráfagas Erróneas (EFB, Erroneous Frame Burst): longitud media de las ráfagas de tramas erróneas a lo largo de la medida.

Longitud Máxima de Ráfagas Erróneas: longitud de la mayor de las ráfagas de tramas erróneas recibidas a lo largo de la medida.

Pérdida MAC Residual: porcentaje de datagramas IP no recuperados por el mecanismo ARQ definido en el estándar IEEE 802.11. Se corresponde con el porcentaje de datagramas IP que se habrían perdido si el FEC no hubiera estado presente.

Para analizar el comportamiento de TCP son necesarios algunos parámetros adicionales. Todos cubren estadísticas específicas de la operación de dicho protocolo y son proporcionadas por las herramientas de análisis empleada a lo largo de la campaña de medidas (*tcpdump y tcptrace*):

Retransmisiones: número total de segmentos de datos TCP retransmitidos por la entidad TCP transmisora.

Máximo número de retransmisiones: número máximo de veces que un mismo segmento de datos se retransmite a lo largo de toda la transferencia.

Tiempo de inactividad máximo: máximo periodo de tiempo durante el que la entidad transmisora no envía ningún segmento.

Reconocimientos triplicados: número de reconocimientos que llegan por tercera vez al transmisor.

Además, es posible estimar la manera en la que se comportó el FEC, calculando las siguientes métricas:

Capacidad correctora: porcentaje de tramas corregidas frente al número total de tramas erróneas y protegidas que llegan al módulo.

Ganancia FEC: porcentaje de datagramas IP que el FEC ha sido capaz de recuperar, frente al número de pérdidas que se habría producido si el FEC no hubiera estado presente.

Con objeto de analizar cómo ha sido la adaptación del módulo a la calidad del canal se recoge también el porcentaje de tramas protegidas recibidas sin error así como el porcentaje de tramas no protegidas recibidas con error.

En cuanto al módulo Snoop, se mide el número de segmentos localmente retransmitidos por el módulo.

Por último hay que decir que todas las estadísticas que se recogen en este artículo se refieren al flujo de datos aunque el módulo FEC también se aplica a los reconocimientos.

5 Optimización Gradual

En este apartado se muestra la evolución en las prestaciones del protocolo TCP a medida que se van añadiendo las técnicas Snoop y FEC a evaluar. Con objeto de proporcionar un análisis más amplio, el MT se ha situado en dos posiciones diferentes denominadas Posición 1 y Posición 2. En este sentido, en la Figura 3 se muestran la distribuciones de la SNR medida en cada una de las posiciones. En el caso de la Posición 1 se incluye, además, la distribución de la tasa de error de trama (FER) por SNR, distribución que se utilizó para fijar los umbrales de 15 y 18 dB que maneja el módulo FEC. Observando dichas figuras se puede pensar que el módulo FEC no estará codificando todas las tramas cuando el MT está situado en la Posición 2 pero sí lo hará cuando está en la Posición 1. Todos los experimentos cuyos resultados se mostrarán a continuación se han obtenido a la velocidad más alta que especifica el estándar IEEE 802.11b, 11 Mbps, velocidad a la que el impacto de los errores es mayor que en el resto de las velocidades permitidas por el estándar.

La Tabla 1 compara los resultados obtenidos en la posición 1 para las cuatro campañas de medidas realizadas: TCP Reno nativo, TCP con Snoop, TCP con FEC y TCP con FEC y Snoop. Mientras que en el caso de un canal libre de errores, el rendimiento que se alcanza con la implementación de TCP Reno es de 5 Mbps [9], éste cae por debajo de los 2.5 Mbps en la Posición 1, caracterizada por presentar una gran

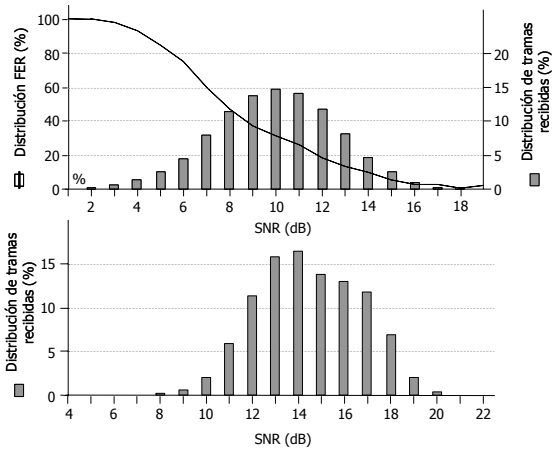


Figura 3. Distribución de SNR observada en las dos posiciones del MT, Posición 1 y Posición 2

Tabla 1. Comparación de los resultados en la Posición 1

#Test	Tput (Mbps)	FER	Pérdida MAC Residual	EFB Media	EFB Max	Retx TCP	#Max Retx	Max idle (sg)	Triple ACK´s	Retx Snoop	Capacidad Corrección	Ganancia FEC	Tramas codif. correct	Tramas erróneas sin codf.
TCP RENO NATIVO														
1	0.55	41.85%	7.2%	2.43	104	620	9	28.16	114					
2	0.67	36.01%	3.5%	2.08	109	264	9	39.68	103					
3	1.19	29.17%	4.2%	2.27	89	314	6	8.32	68					
4	2.39	25.53%	2.9%	2.08	43	217	5	1.28	61					
5	1.30	21.19%	3.9%	3.11	113	321	7	8	39					
SNOOP														
1	1.52	45.17%	7.8%	2.35	155	97	3	1.71	2	620				
2	2.41	35.17%	5.3%	2.29	105	229	3	0.66	3	361				
3	2.09	33.11%	6.0%	2.72	200	218	2	1.7	3	342				
4	2.23	30.55%	5.3%	2.64	171	143	3	1.34	3	386				
5	2.07	30.26%	5.8%	3.07	128	112	3	6	3	377				
FEC														
1	2.01	59.6%	20.9%	5.04	282	68	3	0.54	25		80.60%	95.77%	100%	0%
2	2.56	43.5%	11.2%	4.05	220	137	2	0.29	24		74.69%	84.93%	100%	0%
3	2.67	42.9%	9.2%	3.67	315	50	2	0.43	6		92.14%	94.73%	100%	0%
4	2.64	35.8%	4.6%	2.9	224	114	3	0.44	13		83.83%	76.35%	100%	0%
5	2.55	32.3%	7.3%	4	193	138	3	1.24	19		67.8%	76.11%	100%	0%
FEC + SNOOP														
1	2.82	40.2%	6%	2.3	62	7	1	0.16	3	6	98.35%	97.14%	100%	0%
2	2.37	37.5%	5.6%	2.28	114	66	2	1.98	5	144	40.42%	54.82%	100%	0%
3	2.60	35.6%	6.8%	2.86	119	41	2	0.66	5	197	43.23%	57.41%	100%	0%
4	2.97	33.4%	2.5%	1.71	48	6	1	0.95	3	22	84.06%	85.49%	100%	0%
5	3.24	29%	5.2%	3.05	128	49	1	0.41	3	29	84.58%	79.44%	100%	0%

variabilidad y una alta tasa de errores. Asimismo, la longitud de las ráfagas de tramas erróneas es considerable, entre 50 a 100 tramas consecutivas recibidas con error de CRC, lo que produce sucesivas retransmisiones del mismo segmento TCP (hasta 9 en el peor caso). Otra consecuencia de estas largas ráfagas es el que se alcancen temporizadores de retransmisión de TCP (RTO, Retransmission Time Out) del orden de las decenas de segundos, lo que hace que el rendimiento caiga por debajo de 1 Mbps, como es el caso de las medidas 1 y 2.

Al añadir el agente Snoop se puede observar que, para valores similares o incluso peores, de tasa de error de trama y ráfagas de tramas erróneas, el rendimiento se estabiliza entre 1.5 y 2.5 Mbps, sin alcanzar valores inferiores al Mbps en ningún caso. En particular, al comparar las medidas 1 de TCP Reno y TCP con Snoop, se puede observar que el rendimiento de éste último es tres veces el que se alcanza con la implementación de TCP nativa, para condiciones similares de estado del canal, es decir, pérdida MAC residual algo superior al 7% y ráfagas máximas de tramas erróneas de algo más de 100 tramas, con una ráfaga media de, aproximadamente, 2.5 tramas. Añadir que el agente Snoop hace que se reduzca considerablemente el número de segmentos TCP que retransmite el FH así como el número máximo de retransmisiones del mismo segmento, que no supera las 3 retransmisiones en ningún caso. En esta misma línea se puede observar cómo disminuye el número de reconocimientos triples que alcanzan la entidad TCP transmisora del FH y la duración de los tiempos de inactividad que se producen en dicha entidad. Sin embargo, a pesar de esta notable mejora, el agente Snoop no es capaz de ocultar completamente las pérdidas de paquetes debidas a las imperfecciones del medio inalámbrico, por lo que el FH tiene que retransmitir un número no despreciable de segmentos, aunque mucho menor que en el caso de que el agente no hubiera estado presente.

En contraste con el caso Snoop, en el que no se requiere añadir sobrecarga a los paquetes, cuando se aplica el módulo FEC, el tamaño máximo del segmento TCP (MSS, Maximum Segment Size) se reduce de 1448 a 1385 bytes debido a la redundancia que añade el codificador FEC. Aún así esta técnica también mejora el comportamiento de TCP en esta posición. En el caso de las peores condiciones del canal, el rendimiento alcanzado con el FEC es cinco veces mejor que el que se obtiene con el TCP Reno nativo. La capacidad de corrección alcanzada es, en todos los casos, del orden, o incluso superior, a la considerada durante la fase de diseño del codificador. La ganancia FEC es superior al 76%, por lo que las retransmisiones de segmentos TCP se reducen considerablemente. Adicionalmente se puede observar que el tiempo máximo de inactividad en el transmisor no supera los 2 sg, lo que hace que las ráfagas de tramas erróneas sean más largas que en el caso de no estar activo el módulo FEC. Se puede concluir que la mejora alcanzada con este módulo es sensiblemente superior a la que se obtiene con el módulo Snoop.

Aunque, como se acaba de concluir, el módulo FEC ha resultado ser bastante efectivo a la hora de ocultar los errores y las pérdidas de paquetes a la capa TCP, puede haber situaciones en las que el comportamiento del módulo no sea tan bueno debido a las condiciones particulares del canal inalámbrico. En este sentido y teniendo en cuenta que la carga computacional que

introduce el módulo Snoop no es relevante, puede ser por tanto beneficioso utilizar ambas técnicas simultáneamente. Así, si se analizan los resultados obtenidos en el caso de aplicar los módulos FEC y Snoop de forma simultánea se puede observar que cuando el FEC se comporta de forma eficiente (es decir, cuando la ganancia FEC es cerca del 100%), el agente Snoop no tiene un impacto relevante en el rendimiento final. Por otro lado, en dos de las medidas realizadas, la capacidad de corrección del FEC ha sido baja, por debajo del 50% de las tramas erróneas, y en este caso, el agente Snoop ayuda a mantener el rendimiento en un nivel alto. Un claro ejemplo de esto se encuentra al comparar las medidas 3 de los casos FEC y Snoop con FEC. A pesar de ser peores las condiciones del canal en el segundo caso y de que la capacidad de corrección tampoco ha sido tan alta como en el primer caso, el agente Snoop hace que se alcance el mismo rendimiento en ambos casos. La mejora que supone la combinación de las técnicas FEC y Snoop es aún más significativa cuando se compara con los resultados obtenidos en el caso de aplicar sólo el módulo Snoop. Si se comparan las medidas 5 de ambos casos, en las que las condiciones del canal han resultado ser bastante parecidas, se puede observar que la propuesta Snoop con FEC alcanza 3.2 Mbps frente a 2 Mbps en el caso de aplicar sólo el agente Snoop.

En contraste con los resultados mostrados en la Tabla 1 para el MT situado en la Posición 1, la mejora que se alcanza cuando se combinan ambas técnicas con el MT en la Posición 2 no es tan importante (como se puede observar en la Tabla 2), excepto para aquellos casos en los que la tasa de error de trama observada supera el 10%, en los que el incremento del rendimiento es del 100%. En el resto de las medidas el mecanismo de retransmisión del estándar 802.11, con las tres retransmisiones que realizan las tarjetas inalámbricas utilizadas, es suficiente para contrarrestar la presencia de los errores producidos por el canal inalámbrico. De cualquier modo, la activación de los dos módulos no implica una degradación del rendimiento digna de mencionar.

Por último, es importante destacar que en el caso de que el FEC no se haya adaptado correctamente a la SNR del canal durante el experimento (la adaptación se hace cada 5 sg), el efecto del módulo Snoop cuando se combinan ambas técnicas es más importante que en la Posición 1.

6 Conclusiones

Este artículo muestra, desde un punto de vista puramente experimental, los beneficios de combinar diferentes técnicas de nivel de enlace en el rendimiento del protocolo TCP sobre enlaces IEEE 802.11b con alta probabilidad de sufrir errores a ráfagas. En particular, se ha evaluado el comportamiento del protocolo cuando se añaden un agente Snoop y un esquema FEC, tanto por separado como cuando se aplican de forma simultánea. Cuando las condiciones del enlace son malas, como es el caso de un entorno de oficinas con obstáculos metálicos y personas que se mueven en el canal radio (caracterizado por un valor de SNR medio de 10 dB), el esquema FEC se comporta mejor que el agente Snoop, aunque éste último también mejora de forma considerable el rendimiento. Adicionalmente, se ha observado que el combinar ambas técnicas es especialmente importante en aquellos casos en los que el FEC no alcanza la capacidad de corrección para la que fue diseñado.

Tabla 2. Comparación de los resultados en la Posición 2

#Test	Tput (Mbps)	FER	Pérdida MAC Residual	EFB Media	EFB Max	Retx TCP	#Max Retx	Max idle (sg)	Triple ACK´s	Retx Snoop	Capacidad Corrección	Ganancia FEC	Tramas codif. correct	Tramas erróneas sin codf.
TCP RENO NATIVO														
1	1.56	15.7%	1.8%	2.18	132	131	7	11.2	25					
2	2.34	12.6%	2.0%	2.94	98	157	5	4.8	10					
3	3.5	9.02%	1.5%	3.07	40	120	4	1.68	21					
4	4.36	5.21%	0.03%	1.3	6	1	1	0.2	1					
5	4.84	2.53%	0%	1.18	2	0	0	0.02	0					
SNOOP														
1	3.64	18.12%	2.5%	2.5	92	42	1	0.31	5	236				
2	3.68	15.48%	1.8%	2.21	40	37	1	0.45	5	244				
3	4.03	11.48%	1.5%	2.24	47	39	1	0.54	4	110				
4	4.59	6.44%	0.22%	1.33	15	6	1	0.17	3	73				
5	4.97	0.68%	0.03%	1.35	6	2	1	0.11	1	3				
FEC														
1	3.72	15.9%	2.6%	2.95	141	54	2	0.29	12		76.63%	77.5%	100%	0%
2	3.77	14%	1.2%	1.53	22	47	1	0.07	2		95.91%	48.91%	52.12%	35.46%
3	4.15	8.4%	1.4%	3.35	214	4	1	0.17	1		79.28%	96.23%	100%	0%
4	4.51	7.1%	0.17%	1.34	8	13	1	0.24	7				0%	0%
5	4.78	0.4%	0.03%	2	7	0	0	0.09	0		61.54%	100%	71.1%	18.75%
FEC + SNOOP														
1	3.64	19.3%	4.1%	4.77	150	89	2	0.34	2	71	68.81%	55.52%	39.73%	13.83%
2	4.14	7.4%	1.2%	3.37	64	18	1	0.31	3	29	80.03%	75.27%	100%	0%
3	4.54	3.6%	0.7%	3.53	154	46	2	0.76	1	25	58.74%	15.09%	20.73%	16.26%
4	4.56	1.4%	0.3%	8.37	37	21	1	0.62	1	19	44.55%	22.73%	99.95%	0%
5	4.86	1.4%	0%	1.08	3	0	0	0.1	0	0			0.13%	100%

Cuando las condiciones del canal son mejores (SNR media en torno a 14 dB), la ganancia que se consigue con el agente Snoop es del orden de la que se alcanza con el módulo FEC, siendo dicha mejora significativa cuando la FER observada es superior al 10%. Además, en este caso, combinar ambos métodos tiene aún un mayor efecto que en la otra posición, principalmente debido a que el agente Snoop ayuda a mejorar el comportamiento del protocolo TCP cuando la adaptación del módulo FEC no ha sido correcta. Por otro lado, se ha comprobado que el método de retransmisión del estándar es suficiente para solventar el problema de los errores cuando la FER es baja, aunque se ha demostrado que la sobrecarga que introducen ambos módulos es despreciable.

Es necesario destacar que esta importante mejora en el comportamiento de TCP se consigue a pesar de que no puede existir ninguna interacción entre las dos técnicas propuestas y el mecanismo de retransmisión "cerrado" que utilizan las tarjetas 802.11 empleadas en este trabajo. Es más que probable que se pueda alcanzar un rendimiento superior si fuese posible una interacción real entre los distintos mecanismos. Teniendo en cuenta que están apareciendo productos comerciales más abiertos, se acometerá en el futuro la integración de las técnicas FEC y Snoop con esas nuevas interfaces.

Referencias

[1] B. Liu, D. L. Goeckel, and D. Towsley. "TCP-cognizant adaptive forward error correction in wireless networks", in Proc. GLOBECOM, Taipei, Taiwan, Nov. 2002, pp. 2128-2132.

[2] G. Yang, R. Wang, F. Wang, M. Y. Sanadidi and M. Gerla, "TCPW with Bulk Repeat in Next Generation Wireless Networks," in Proc. ICC, Anchorage, Alaska, May 2003, pp 674-678.

[3] D.C. Feldmeier, A.J. McAuley, J.M. Smith, D.S. Bakin, W.S. Marcus and T.M. Raleigh, "Protocol Boosters," in IEEE Journal on Selected Areas in Communications, vol. 16, no 3, pp 437-444, April 1998.

[4] M. Zorzi, A. Chockaligam and V. Tralli, "Wireless TCP Performance with Link Layer FEC/ARQ", in Proc. ICC, Vancouver, Canada, June 1999, pp.1212-1216.

[5] D. Gibson, "Wireless Networking with Linux and IEEE 802.11b," in Proc. 8th Int. Linux-Kongress, Enschede, The Netherlands, Nov. 2001.

[6] H. Balakrishnan, S. Seshan and R. H. Katz, "Improving Reliable Transport and Handoff Performance in Cellular Wireless Networks", ACM Wireless Networks, vol. 1, no 4, pp. 469-482, Dec. 1995.

[7] H. Balakrishnan, V. Padmanabhan, S. Seshan and R. Katz, "A Comparison of Mechanisms for Improving TCP Performance over Wireless Links," IEEE/ACM Trans. on Networking, vol 5, no 6, pp. 756-769, 1997.

[8] P. Mähönen, N. Passas, G. Orphanos, L. Muñoz, A. Marshall, D. Melpignano, T. Inzerilli, F. Lucas, M.Vitiello, M. García and T. Saarinen, "Platform-Independent IP Transmission over Wireless Networks: The WINE Approach," IEEE Personal Comm. Mag., vol. 8, no 6, pp. 32-40, Dec. 2001.

[9] L. Muñoz, M. García, J. Choque, R. Agüero and P. Mähönen, "Optimizing Internet Flows over IEEE 802.11b Wireless Local Area Networks: A Performance Enhancing Proxy Based on Forward Error Correction", IEEE Comm. Mag., vol 39, no 12, pp 60-67, Dec. 2001.

Planificador GPS con desacoplamiento de ancho de banda y retardo

José Ramón Piney, Sebastià Sallent
Departamento de Ingeniería Telemática. Universidad Politécnica de Cataluña
Barcelona
E-mail: {jpiney, sallent}@mat.upc.es

Abstract *The possibility of providing independent guarantees of bandwidth and delay is an important aspect to be considered in the services disciplines at the switching nodes. We propose a modified version of Generalized Processor Sharing scheduler with weights that are variable in time and a method that permits to calculate these weights. Besides we show that service curve offered by this scheduler guarantee the delay bound and long-range rate of input traffic.*

1. Introducción

La elección de un planificador adecuado en un conmutador es uno de los aspectos más importantes para proporcionar garantías de servicio en redes de alta velocidad. Éste determina como interactuan los paquetes de diferentes conexiones entre sí. Las garantías de servicio de cada conexión son independientes y distintas unas de otras. Esto significa que una aplicación puede ser muy sensible al retardo pero que necesite un ancho de banda pequeño, mientras que otra puede necesitar una gran ancho de banda pero funcionar adecuadamente con un retardo amplio. Sin embargo en muchas de las disciplinas que existen (Weighted Fair Queueing-WFQ, Self Clock Fair Queueing-SCFQ, Virtual Clock-VC)[1], el retardo se controla mediante una cota de ancho de banda asignado. Puesto que el retardo y el ancho de banda no se pueden modificar independientemente, se puede desperdiciar ancho de banda ya que en algunos casos para asegurar un retardo pequeño es necesario realizar una gran reserva de ancho de banda aún en el caso que la aplicación requiera un ancho de banda pequeño. También hay otras disciplinas de servicio que aseguran un retardo, pero no controlan el ancho de banda que consume la aplicación (como en el caso de Earliest Deadline First-EDF) [1] lo cual puede producir que una conexión mal intencionada acapare el ancho de banda del enlace de salida.

En este trabajo se presenta un planificador ideal basado en el Procesador Generalizado Compartido (GPS por sus siglas en ingles) [2] que evita la dependencia directa entre ancho de banda asignado y retardo y que proporciona una curva de servicio no lineal [8] permitiendo especificar ambos parámetros de forma separada. Además se muestra un esquema para el cálculo de dichas funciones, después de realizar un control de admisión.

[1]En este trabajo se utilizará indistintamente sesión o flujo.

2. Planificador GPS

De acuerdo con la descripción de [1] y [2], un planificador GPS es un servidor conservativo, (es decir que mientras haya paquetes en espera, el servidor estará ocupado) que opera a una tasa fija r, y está caracterizado por un conjunto de números reales positivos $\phi_1, \phi_2, ..., \phi_N$. Si denotamos $S_i(\tau, t)$ como la cantidad de tráfico de la sesión [1] i servido en un intervalo $(\tau, t]$, entonces el servidor GPS se define como uno en el cual

$$\frac{S_i(\tau,t)}{S_j(\tau,t)} \geq \frac{\phi_i}{\phi_j},\ j = 1, 2, ..., N$$

para cualquier sesión i que está continuamente con paquetes pendientes (en espera en la cola) en el intervalo $(\tau, t]$.

Sumando sobre todas las sesiones j:

$$S_i(\tau,t)\sum_j \phi_j \geq \phi_i r(t-\tau)$$

$$S_i(\tau,t) \geq \frac{\phi_i}{\sum_j \phi_j} r(t-\tau)$$

donde $(t-\tau)r = \Sigma_j S_j(t-\tau)$, con lo cual a la sesión i se le garantiza una tasa mínima de

$$g_i = \frac{\phi_i}{\sum_j \phi_j} r$$

Por lo tanto, mientras la tasa asignada a cada sesión i, r_i cumpla $r_i \leq g_i$, la sesión i tendrá garantizada dicha tasa indepedientemente de las demandas de las demás sesiones.

Si el tráfico de una sesión i que llega al servidor está restringido por un *leaky-bucket* con parámetros (σ_i, ρ_i), el retardo del flujo i, d_i, introducido por el servidor GPS está acotado por [3]

$$d_i \leq \frac{\sigma_i}{g_i}$$

como se puede observar existe una relación inversamente proporcional entre ancho de banda y retardo.

3. Planificador GPS con desacoplamiento de ancho de banda y retardo

Con la intención de que el retardo no dependa directamente del ancho de banda, se considera la posibilidad de variar los pesos ϕ_i's para tener la flexibilidad de tratar las sesiones según las restricciones que necesitamos que cada sesión cumpla.

Si hacemos variables los pesos y suponemos sin pérdida de generalidad que $\tau = 0$, el tráfico servido de la sesión i será

$$S_i(0,t) \geq \int \frac{\phi_i(t)}{\sum_j \phi_j(t)} r dt$$

Si

$$f_i(t) = \int \frac{\phi_i(t)}{\sum_j \phi_j(t)} r dt \quad (1)$$

$$S_i(0,t) \geq f_i(t) \quad (2)$$

donde $f_i(t)$ se puede considerar como una curva de servicio [4, 5, 6], la cual es funcion de los pesos o en este caso de las funciones de peso $\phi_i(t)'s$.

3.1. Curva de servicio de un planificador GPS desacoplado

En esta sección presentamos el esquema para obtener una curva de servicio que cumpla con unas restricciones de tasa a largo plazo y retardo y veremos como no siempre es posible cumplir estas condiciones con un servidor GPS. Suponemos una sesión de entrada i regulada, es decir, existe una función creciente en sentido amplio α_i tal que [7]

$$A_i(t) - A_i(s) \leq \alpha_i(t-s), \text{ para cualquier } s \leq t$$

donde $A_i(t)$ es el tráfico de entrada de la sesión i en el intervalo $(0,t]$. En particular, consideramos la entrada regulada por un leaky-bucket; $\alpha_i(t) = \rho_i t + \sigma_i$. Además suponemos que $E[A_i(t) - A_i(s)] \leq \rho_i(t-s)$, para cualquier $s \leq t$, donde

$$\rho_i = \lim_{t\to\infty} \frac{\alpha_i(t)}{t} = \lim_{t\to\infty} \alpha_i'(t)$$

que representa la tasa a largo plazo. Basándonos en lo anterior, nos interesa que la curva de servicio $f_i(t)$ que ofrece el scheduler garantice la tasa a largo plazo de la sesión i, ρ_i. Si además ponemos la restricción de que $f_i(t)$ sea lineal en el intervalo $(0,\infty)$ (condición que impone el servidor GPS) tenemos que

$$\lim_{t\to\infty} \frac{f_i(t)}{t} = f_i'(t) \text{ para toda t}$$

por lo tanto

$$f_i'(t) = \rho_i$$

con lo cual tenemos que

$$f_i(t) = \int_0^t \rho_i dt = \rho_i t + C$$

En este caso el valor de C se obtiene de la restricción de que la cota de retardo máximo de la sesión i debe ser d_i.

El retardo máximo D es la desviación horizontal $h(\alpha,\beta)$, introducida en [5], y definida como

$$h(\alpha,\beta) = \sup_{t\geq 0}\{\inf\{d \geq 0 \text{ tal que } \alpha(t) \leq \beta(t+d)\}\}$$

donde α y β son dos funciones crecientes en sentido amplio.

Entonces suponiendo un flujo restringido por una curva de llegada α que atraviesa un sistema que ofrece una curva de servicio β, el retardo virtual $d(t)$ para toda t satisface:

$$d(t) \leq h(\alpha,\beta).$$

En nuestro caso, $\alpha(t) = \rho_i t + \sigma_i$ y $\beta(t) = \rho_i t + C$ y $d = d_i$ por lo tanto, si queremos que

$$\alpha(t) \leq \beta(t+d_i) \quad (3)$$

Necesitamos que C cumpla

$$C \geq \sigma_i - \rho_i d_i$$

con lo cual la curva de servicio debe cumplir

$$f_i(t) \geq \rho_i t + \sigma_i - \rho_i d_i$$

y

$$S_i(0,t) \geq \rho_i(t-d_i) + \sigma_i$$

Si $\sigma_i = \rho_i d_i$, esto significa que con la tasa asignada a largo plazo se cumple también con la cota de retardo por lo que la curva de servicio queda como una función lineal que pasa por el origen,

$$f_i(t) \geq \rho_i t$$

en este caso, de (1) tenemos,

$$\rho_i = \frac{\phi_i(t)}{\sum_j \phi_j(t)} r$$

esto se puede lograr con pesos constantes, con lo que correspondería a un sistema GPS.

Si $\sigma_i < \rho_i d_i$ entonces la curva de servicio lineal con pendiente ρ_i puede iniciarse en $t = d_i - \frac{\sigma_i}{\rho_i}$. Esto indica dos cosas, en el caso que se considerara

iniciar el servicio un intervalo de tiempo $d_i - \frac{\sigma_i}{\rho_i}$ después de la llegada del primer paquete de cada periodo de ocupación de la sesión i, el sistema se consideraría no conservativo y si el servicio se iniciara en el instante en que llega el paquete, significaría que estaría desperdiciando servicio al asegurar un retardo mucho menor a la cota necesaria. En la Fig. 4 se muestran ambos casos. Esto concuerda con el trabajo de Schmitt [8], donde se indica que el uso de una curva lineal lleva al desperdicio de los recursos asignados, en particular para flujos de tipo "bajo ancho de banda, retardo pequeño".

Finalmente, si $\sigma_i > \rho_i d_i$, esto significa que en $t = 0$ ya se deben transmitir $\sigma_i - \rho_i d_i$ bits, lo cual no es posible pues se requiriría una tasa infinita.

Para este último caso, se podría considerar una función lineal a trozos, donde la primera parte tiene una tasa r_i ($r_i \leq r$)que garantiza la cota de retardo d_i y después una parte que permite garantizar una tasa ρ_i a largo plazo, como se indica en la Fig. 5 (donde $\sigma_i > \rho_i d_i$).

Para cumplir con la cota de retardo máximo,

$$f_i(d_i) = r_i d_i = \sigma_i$$

por lo tanto se requiere una tasa mínima de r_i.

Finalmente la curva de servicio $f_i(t)$ quedaría como

$$f_i(t) = \begin{cases} r_i t & t \leq d_i \\ \rho_i(t - d_i) + \sigma_i & t > d_i \end{cases} \quad (4)$$

Así $f_i(t)$ es una posible función que permite asegurar una determinada tasa a largo plazo y una cota de retardo máximo, sin embargo, nos falta encontrar la función de pesos $\phi_i(t)$ que nos de $f_i(t)$. Si la función de pesos es lineal por trozos se cumple

$$f_i(t) = \frac{\phi_i(t)}{\sum_j \phi_j(t)} rt \quad (5)$$

Lo más simple para obtener la función de pesos, es suponer que $\sum_j \phi_j(t) = 1$ para toda t con lo cual

$$f_i(t) = \phi_i(t) rt.$$

Además a partir de la ec. 2 y como $rt = \Sigma_j S_j(t)$, las curvas de servicio $f_i(t)$ deben cumplir

$$\sum_i f_i(t) \leq rt \quad (6)$$

en caso de que esto no sea así, significaría que no se pueden cumplir con las restricciones de todas las sesiones.

Finalmente para la curva de servicio representada por (4) , una posible función de pesos $\phi_i(t)$ sería

$$\phi_i(t) = \begin{cases} 0 & 0 < t \\ \frac{r_i}{r} & 0 \leq t < d_i \\ \frac{\rho_i}{r} & d_i \leq t \end{cases}$$

4. Control de admisión

Antes de obtener la curva de servicio de una nueva sesión, se debe decidir si dicha sesión puede ser aceptada. En [9] aparece un esquema de control de admisión para un planificador EDF (Earliest Deadline First), en donde se considera un sistema con un conjunto de flujos con cotas de retardo máximo que cumplen $d_1 \leq d_2 \leq ... \leq d_N$ y se define

$$F(t) = rt - \sum_{i \in N} A_i(t - d_i) \quad (7)$$

Si $F(t) \geq 0$ para toda t significa que se pueden cumplir las cotas de retardo de todos los flujos.

EDF es un planificador óptimo en cuanto a retardo, puesto que minimiza el retardo máximo de todos los paquetes. Por lo tanto, si $F(t) \geq 0$ para toda t no se cumple, esto significa que no es posible encontrar funciones de servicio $f_i(t)$ que garanticen a cada flujo i un retardo d_i, pues de (3) tenemos que $f_i(t)$ debe cumplir

$$A_i(t - d_i) \leq f_i(t)$$

Para el caso en el que cada flujo esté acotado por un leaky-bucket con parámetros (σ_i, ρ_i), tenemos que la ec. 7 queda como

$$F(t) = rt - \sum_{i \in N} (\sigma_i + \rho_i(t - d_i)) U(t - d_i)$$

donde $U(t) = 1$ si $t \geq 0$ y 0 en otra parte.

Lo anterior se puede convertir en:

$$F(d_j) = rd_j - \sum_{i \leq j} \sigma_i + \rho_i(d_j - d_i)$$

con lo que las garantías de ancho de banda y retardo máximo se cumplen si $F(d_j) \geq 0$ para toda $j \in N$.

Si un nuevo flujo acotado por un leaky-bucket con parámetros (b_f, ρ_f) y con una cota de retardo máxima d_f, se quiere agregar a un conjunto de N flujos donde la cota de retardo del nuevo flujo es tal que $d_{b-1} < d_f \leq d_b$, se debe cumplir que:

1)

$$\sum_{i=1}^{N+1} \rho_i \leq r$$

2)

$$F(d_j) = rd_j - \sum_{i \leq j} \sigma_i + \rho_i(d_j - di) \geq 0 \quad (8)$$

para toda $i \in \{1, 2, ..., N+1\}$ y para toda $j \in \{b, b+1, ..., N+1\}$. Como el nuevo flujo se ha insertado a partir de $b-1$, esta condición solo se debe probar a partir de $j = b$.

4.1. Cálculo de las funciones de servicio

Una vez se ha aceptado un flujo, hay que calcular la función de servicio que permite cumplir con sus restricciones de tasa a largo plazo y retardo máximo, pero además tomando en cuenta no afectar las restricciones del resto de flujos.

Seguimos suponiendo que cada flujo i está caracterizado por una envolvente $(\sigma_i,\ \rho_i)$, y además los flujos están ordenados de forma que las cotas de retardo máximo cumplen: $d_1 \leq d_2 \leq d_3 \leq ... \leq d_m \leq ... \leq d_{n-1} \leq d_n$.

Para el cálculo de una función de servicio para el flujo m primero se obtendrá una función de peso ϕ_m con la cual,

$$f_m(t) = \int_0^t \phi_m(t)dt$$

Para aplicar el scheduler GPS con desacoplamiento, realmente no es necesario calcular la función de servicio, sino que con la función de pesos es suficiente.

La función de peso tendrá la siguiente forma:

$$\phi_m(t) = \begin{cases} 0 & t < 0 \\ \phi_m^1 & 0 \leq t < d_1 \\ \phi_m^2 & d_1 \leq t < d_2 \\ . & \\ . & \\ . & \\ \phi_m^i & d_{i-1} \leq t < d_i \\ . & \\ . & \\ . & \\ \phi_m^m & d_{m-1} \leq t < d_m \\ \phi_m^{m+1} & t \geq d_m \end{cases} \quad (9)$$

donde ϕ_i^j representa el j-ésimo trozo de la función de peso del flujo i. En la Fig. 6 se puede ver una función de servicio obtenida a partir de una función de peso.

Como se puede observar, la función de peso de un flujo i (donde i es la posición del flujo dentro del conjunto de los n flujos) tendrá $i+1$ trozos lineales, donde la tasa del último trozo corresponde a la tasa a largo plazo del flujo, $\phi_m^{m+1} = \rho_m$. De esta forma,

$$\text{tasa a largo plazo } = \lim_{t\to\infty} \frac{f_m(t)}{t}$$

$$\text{tasa a largo plazo } = \lim_{t\to\infty} \frac{1}{t} \int_0^t \phi_m(t)dt$$

$$= \lim_{t\to\infty} \frac{1}{t}(\int_0^{d_1} \phi_m(t)dt + ... + \int_{d_m}^t \phi_m(t)dt)$$

como $\phi_m(t)$ es lineal por trozos, obtenemos

$$\text{tasa a largo plazo } = \phi_m^{m+1} +$$

$$\lim_{t\to\infty} \frac{1}{t}(\phi_m^1 d_1 ... + \phi_m^m d_{m-1} - \phi_m^m d_{m-1})$$

Finalmente

$$\text{tasa a largo plazo } = \phi_m^{m+1} = \rho_m$$

Se puede comprobar que para obtener el j-ésimo trozo ($j \leq m$ y $i < n$) de la función de pesos de un flujo m, primero se debe calcular,

$$\phi_m^j = \frac{\sigma_m - \sum_{i=1}^{j-1}(d_i - d_{i-1})\phi_m^i}{d_m - d_{j-1}} \text{ donde } d_0 = 0$$

después se debe comprobar si este peso no afecta a los pesos de los flujos que se encuentran por delante de i, es decir que tienen una cota de retardo más pequeña.

Así si

$$\phi_m^j \leq 1 - \sum_{i=1}^{j-1} \rho_i - \sum_{i=j}^{m-1} \phi_i^j \quad (10)$$

y entonces se puede concluir que los restantes pesos tendrán este mismo valor

$$\phi_m^j = \phi_m^{j+1} = ... = \phi_m^{m-1} = \phi_m^m \quad (11)$$

en caso contrario

$$\phi_m^j = 1 - \sum_{i=1}^{j-1} \rho_i - \sum_{i=j}^{m-1} \phi_i^j \quad (12)$$

y se tendrá que calcular el siguiente trozo $j+1$.

Y el último trozo, como ya se indicó es,

$$\phi_m^{m+1} = \rho_m$$

Finalmente, el último flujo se queda con el peso sobrante, por lo que en este caso,

$$\phi_n^j = 1 - \sum_{i=1}^{j-1} \rho_i - \sum_{i=j}^{n-1} \phi_i^j$$

Para ver si la función de servicio obtenida cumple con el retardo, tenemos que el retardo máximo ocurre cuando el flujo transmite un bucket completo. Si suponemos que el flujo m ($m < n$) se pone activo en $t = 0$ entonces en este instante se debe cumplir la ec. (3)

$$A_m(0) \leq f_m(d_m)$$

$$f_m(d_m) = \int_0^{d_m} \phi_m dt$$

$$f_m(d_m) = \sum_{j=1}^{m} \phi_m^j (d_j - d_{j-1}) \quad (13)$$

Además consideramos que hasta el l-ésimo trozo de la función de peso se cumple la ec. 12 y a partir del trozo $l+1$ de la función de pesos se cumplen las ecs. 10 y 11. Entonces

$$f_m(d_m) = \sum_{j=1}^{l} \phi_m^j (d_j - d_{j-1}) + \sum_{j=l+1}^{m} \phi_m^j (d_j - d_{j-1})$$

$$f_m(d_m) = \sum_{j=1}^{l} \phi_m^j (d_j - d_{j-1})$$

$$+ \left(\frac{\sigma_m - \sum_{i=1}^{l} \phi_m^i (d_i - d_{i-1})}{d_m - d_l} \right) \sum_{j=l+1}^{m} (d_j - d_{j-1})$$

Como $\sum_{j=l+1}^{m} (d_j - d_{j-1}) = d_m - d_l$

$$f_m(d_m) = \sigma_m \tag{14}$$

y como $A_m(0) = \sigma_m$, por lo tanto se cumple con el retardo.

Para el caso del flujo final que se queda con el ancho de banda sobrante

$$f_n(d_n) = \sum_{j=1}^{n} \left(1 - \sum_{i=1}^{j-1} \rho_i - \sum_{i=j}^{n-1} \phi_i^j \right) (d_j - d_{j-1})$$

utilizando las ec. 13 y 14 obtenemos

$$f_n(d_n) = d_n - \sum_{i=1}^{n-1} (\sigma_i + \rho_i (d_m - d_i))$$

Finalmente como los flujos han pasado el control de admisión eso significa que la función $F(d_n)$ de la ec. 8 es mayor que cero y por lo tanto

$$f_n(d_n) = \sigma_n$$

Retomando el caso de la sección anterior del flujo aceptado con un leaky-bucket (b_f, ρ_f) y con cota máxima de retardo d_f, el siguiente paso es el cálculo de la función de servicio y además recalcular las funciones de servicio de los flujos $b+1, b+2, ..., n$, que se ven afectados por el nuevo flujo aceptado. En el caso que el flujo nuevo quedara en la última posición $(n+1)$, también se tendría que recalcular la función del flujo n.

5. Consideraciones sobre el GPS desacoplado

Las curvas de servicio que se calculan según el método de la subsección 4.1, consideran el peor caso, esto es que todos los flujos se inician en el mismo instante y que se mantienen continuamente activos.

Para el caso en que un flujo deja de transmitir durante un tiempo y vuelve a estar activo (inicio de un periodo de ocupació de flujo), no se puede tomar el tiempo de ocupación del sistema como tiempo de la función de servicio también, porque no se podría garantizar el retardo de este flujo. Una posibilidad sería calcular un nuevo conjunto de funciones de servicio para ese instante, con condiciones menos restrictivas, pues el tráfico de los flujos que permanecieron activos estará acotado con unos parámetros (σ', ρ), donde $\sigma' \leq \sigma$ (es decir la ráfaga que se puede recibir será menor que en el caso inicial). Sin embargo, esto es complicado. Una solución más simple, es que cuando esto ocurra se inicialice el tiempo de las funciones de servicio pero teniendo en cuenta la ráfaga máxima que el flujo que se ha vuelto activo puede transmitir.

Si suponemos t_u el instante en que la cola del flujo i se vacia por primera vez medido desde la llegada del primer paquete, tenemos

$$A_i(0, t_u) - S_i(0, t_u) = 0$$

Si el paquete que inicia el nuevo período de ocupación del flujo i llega en el instante t_a, entonces el tamaño máximo de la ráfaga que se puede recibir es,

$$\sigma_i' = \text{mín}\{\sigma_i + \rho_i t_a - S_i(t_u), \sigma_i\}$$

Esto significa que a partir de t_a el tráfico estará acotado por

$$A_i(0, t) \leq \sigma_i' + \rho_i t$$

y en el caso que $\sigma_i' < \sigma_i$, se requiere menos tiempo para vaciar el bucket que es lo que determina el retardo máximo. Por lo tanto en vez de iniciar las funciones de servicio desde el origen se pueden iniciar en $(f_i^{-1}(\sigma_i - \sigma_i'), \sigma_i - \sigma_i')$ para el flujo i y para el resto de flujos donde $j \neq i$ en $(f_i^{-1}(\sigma_i - \sigma_i'), f_j(f_i^{-1}(\sigma_i - \sigma_i')))$.

Por otra parte en la aplicación del GPS desacoplado pueden aparecer problemas en casos extremos como el siguiente, donde se ajusta la cota de retardo al mínimo posible y con la restricció $d_1 \leq d_2 \leq d_3$, para las siguientes sesiones:

S	σ	ρ	d
1	25	0,5	25
2	75	0,3	175
3	100	0,2	675

En este caso extremo la única forma de lograr las restricciones anteriores, es que las funciones de peso de cada sesión tengan la siguiente forma:

$$\phi_1(t) = \begin{cases} 1 & 0 \leq t \leq 25 \\ \frac{1}{2} & t \geq 25 \end{cases}$$

$$\phi_2(t) = \begin{cases} 0 & 0 \leq t \leq 25 \\ \frac{1}{2} & 25 \leq t \leq 175 \\ \frac{3}{10} & t \geq 175 \end{cases}$$

$$\phi_3(t) = \begin{cases} 0 & 0 \leq t \leq 175 \\ \frac{2}{10} & t \geq 175 \end{cases}$$

Se puede comprobar que $\sum_{i=1}^{3} \phi_i(t) = 1$. Sin embargo aquí el problema aparece por que durante cierto tiempo algunos pesos son cero.

Si todos los flujos están activos, entonces no hay problema, pues el que algunos tengan una tasa cero $(\phi = 0)$ significa que de momento no deben recibir servicio. Pero que pasa si en un inicio de periodo de ocupación t_0 no está activo el flujo 1.

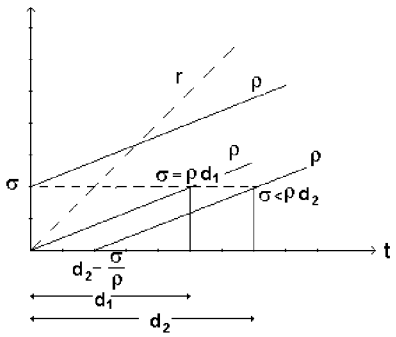


Figura 4: Ejemplo de curva de llegada y de posibles funciones de servicio que aseguran diferentes retardos.

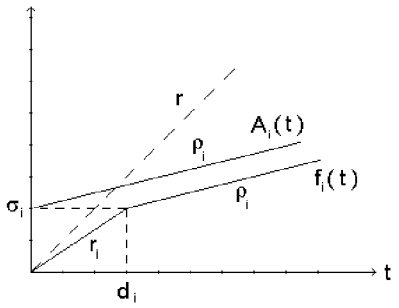


Figura 5: Ejemplo de curva de llegada y curva de servicio lineal por partes.

El servicio sobrante no se reparte entre los flujos activos en ese momento pues los pesos $\phi_2(t_0) = 0$ y $\phi_3(t_0) = 0$, y con ello

$$g_i = \frac{\phi_i}{\sum_j \phi_j} r = 0$$

con lo cual el sistema sería no conservativo. Una posible solución es desplazar la variable de tiempo de forma que coincida con el tiempo en que el peso del flujo en activo con la cota de retardo más grande se hace mayor a cero, es decir según el ejemplo que $t_0 = 175$, de esta forma se tiene la seguridad de que todos los flujos activos en ese momento reciben servicio.

6. Conclusiones y trabajo futuro

En este trabajo se propone un planificador que permite asegurar una determinada función de servicio, la cual permite garantizar una asignación de ancho de banda y una cota máxima de retardo. Para ello, se ha modificado el planificador GPS, haciendo que los pesos varien en función del tiempo. Estos pesos son los que establecen la función de servicio ofrecida por el planificador.

Además se presenta un esquema de control de admisión junto con un método para calcular la función de servicio de un nuevo flujo que haya sido aceptado.

Finalmente, se hacen algunas consideraciones sobre ciertos problemas que puede presentar este planificador y como solucionarlos.

Se debe mencionar que como trabajo futuro se está desarrollando una versión realizable de este planificador ideal.

Agradecimientos

Este trabajo ha sido financiado parcialmente por la CICYT TIC2003-09042-C03-02.

Referencias

[1] H. Zhang, "Service disciplines for guaranteed performance service in packet-switching networks," *Proceedings of the IEEE*, 83(10), pp. 1374–1396, Oct. 1995.

[2] A. K. Parekh and R. Gallager, "A generalized processor sharing approach to flow control in integrated services networks: the single-node case," *IEEE/ACM Trans. on Networking*, 1(3), pp. 334–357, June 1993.

[3] P. Goyal and H. Vin, "Generalized guaranteed rate scheduling algorithms: a framework," *Technical Report TR-95-30*, University of Texas, Austin, September 1995.

[4] R. L. Cruz, "A calculus for network delay, Part I: network elements in isolation," *IEEE Trans. on Information Theory*, 37(1), pp. 114–131, Jan. 1991.

[5] J. Le Boudec and P. Thiran, *Network Calculus.* Springer Verlag, 2002

[6] Cheng-Shang Chang, *Performance Guarantees in Communtication Networks.* Springer, 2000

[7] M. Vojnović and J.-Y Le Boudec, "Stochastic analysis of some expedited forwarding networks ," *Proceedings of INFOCOM '02*, pp. 1004–1013, 2002.

[8] J. Schmitt, "Optimal network service curves under bandwidth-delay decoupling," *Electronics Letters*, 38(6), pp. 297–298, 2002.

[9] V. Firoiu, J. Kurose, and D. Towsley, "Efficient admission control for EDF schedulers," *IEEE* , 43(5), pp. 1518–1535, 1997.

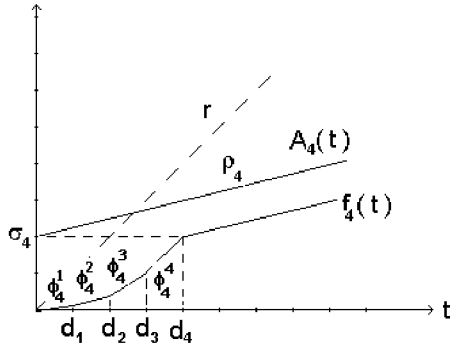


Figura 6: Ejemplo de función de servicio para un flujo con retardo máximo d_4 y tasa a largo plazo ρ.

Aproximación de Árboles Multicast Óptimos en Redes Ad hoc Inalámbricas

Pedro M. Ruiz, Antonio F. Gómez Skarmeta
Depto. Ingeniería de la Información y las Comunicaciones
Fac. Informática, Campus de Espinardo S/N.
30071 - Espinardo (Murcia)
Telefono: 968 36 46 07 Fax: 968 36 41 51
E-mail: {pedrom,skarmeta}@dif.um.es

Abstract *The problem of computing minimum cost multicast trees is known as the Steiner tree problem. This NP-complete problem has been widely studied in wired networks. In this paper, we show that a Steiner tree does not offer the minimum bandwidth multicast tree for wireless multihop networks. The problem needs to be reformulated in terms of minimizing the number of transmissions which are required to deliver a packet from a multicast source to a set of receivers. We show that the new problem is also NP-complete, and we propose some heuristics to approximate such optimal trees. Our simulation results show that the proposed heuristics outperform Steiner trees over a variety of scenarios.*

1. Introducción

Una red ad hoc está formada por un conjunto de nodos, equipados con interfaces inalámbricas. Los nodos que no están lo suficientemente cerca como para comunicarse directamente, usan a otros nodos ad hoc como encaminadores intermedios. Cuando los nodos tienen libertad de movimiento, estas redes suelen denominarse "redes ad hoc móviles". En estas redes móviles no tiene sentido optimizar los árboles multicast, ya que la movilidad haría que dejasen de ser óptimos muy rápidamente. Es por ello que en este artículo nos centramos en sus homólogas estáticas, también conocidas como "mesh networks". Esta última variante está recibiendo una gran atención de la comunidad científica y se consideran una opción simple y barata para que los operadores puedan extender su cobertura para reaccionar rápidamente a demandas temporales.

Por otro lado, IP multicast es una de las áreas que se espera jueguen un papel fundamental en las redes móviles e inalámbricas futuras. De hecho, muchos de los servicios que se preveen en el futuro requieren un importante ancho de banda, y están fuertemente basados en interacciones muchos a muchos. Para soportar este tipo de servicios sobre redes ad hoc inalámbricas se requiere un soporte eficiente de comunicaciones multipunto, que permita salvar las limitaciones de ancho de banda de este tipo de redes.

El problema de la distribución eficiente de tráfico entre un conjuto de emisores y receptores en una red basada en datagramas fue estudiado por Deering [1] a finales de los 80. Se han propuesto varios protocolos de encaminamiento multicast para redes IP fijas tales como DVMRP [2], MOSPF [3], CBT [4] y PIM [5]. Sin embargo, estos protocolos no funcionan bien en redes ad hoc debido a su fragilidad a la hora de soportar la movilidad de los nodos. Aunque podría parecer que estos protocolos son apropiados para redes ad hoc estáticas, la realidad es que no están preparados para estos escenarios inalámbricos. Eso les hace ofrecer soluciones subóptimas, sin ser capaces de aprovechar la naturaleza de multidifusión que ofrece el medio inalámbrico, sin permitir una reducción del consumo de energía de los nodos, etc.

El cálculo del árbol multicast de mínimo costo es conocido como el "Steiner tree". Karp [7] demostró por transformación a la cobertura exacta por tres conjutos que este problema es NP-completo incluso cuando todos los enlaces tienen el mismo coste. Existen algoritmos heurísticos [8] para aproximar este tipo de árboles. Por ejemplo, el algoritmo MST ([9, 10]) ofrece un ratio de aproximación de 2, mientras que Zelikovsky [11] propuso un algoritmo con ratio de aproximación de 11/6. Sin embargo, dado que es muy complicado encontrar heurísticas distribuidas con un bajo coste computacional, la mayoría de los protocolos de encaminamiento multicast existentes se basan en árboles de camino más corto (Shostest Path Tree) o árboles compartidos, que pese a estar muy alejados del óptimo, se pueden calcular muy fácilmente en un tiempo polinomial.

Por los mismos motivos, los algoritmos de encaminamiento multicast para redes ad hoc que se encuentran en la literatura [6] no usan árboles de distribución de mínimo coste. De hecho, para redes inalámbricas la mayor parte de los trabajos de investigación en materia de la reducción de la sobrecarga de datos se han centrado en el caso particular de la inundación ("broad-

cast storm problem"). Sólo algunos autores como Lim y Kim [13] analizaron el problema de minimizar los árboles multicast en redes ad hoc, aunque las heurísticas que presentaron basadas en el MCDS (Minimum Connected Dominating Set) son sólo validas para el problema de la inundación.

Siempre se ha asumido que el árbol multicast con un menor consumo de ancho de banda era el Steiner tree. Esto es cierto en redes fijas, sin embargo, en este artículo demostramos que esto no es cierto en redes ad hoc, en las que la naturaleza broadcast de los enlaces permite reducir el número de transmisiones necesarias para enviar un datagrama a múltiples destinos. El problema de encontrar el árbol multicast de coste mínimo en redes ad hoc ha de reformularse en términos de la minimización del número de retransmisiones de datos necesarias. Las formulaciones existentes asignaban un coste a cada enlace del grafo, y encontraban el árbol multicast cuya suma de los costes de los enlaces era mínimo. De esa forma, asumían de forma implícita que dado un nodo v, éste requiere k transmisiones para enviar un mensaje un número k de vecinos en el árbol multicast. Sin embargo, en un medio broadcast, se puede transmitir ese mensaje a cualquier número de vecinos con una sola transmisión. Es por ello, que el árbol multicast mínimo en una red ad hoc es aquel que conecta la fuente y los receptores requiriendo un número mínimo de transmisiones.

En este artículo mostramos que el Steiner tree no ofrece una solución óptima al problema. Como contribución adicional demostramos que calcular el árbol multicast de coste mínimo es también NP-completo y ofrecemos algoritmos heurísticos para aproximar dichos árboles óptimos que denominamos árboles de mínima sobrecarga de datos. Nuestras simulaciones muestran que los algoritmos propuestos ofrecen mejores resultados que la heurística MST para Steiner trees en una gran variedad de escenarios. Además, se aprecia una enorme reducción en cuanto a coste en comparación con los algortimos de encaminamiento multicast para redes ad hoc actuales basados en "Shortest path trees".

El resto del artículo se organiza como sigue: la sección 2 describe nuestro modelado de la red, nuestra formulación del problema y demuestra que dicho problema es NP-completo. En la sección 3 describimos los algoritmos propuestos. Los resultados de las simulaciones se presentan en la sección 4, y por último, en la sección 5 ofrecemos las conclusiones.

2. Modelo de red y formulación del problema

2.1. Modelo de red

Representamos una red ad hoc como un grafo no dirigido $G = (V, E)$ en el que V es el conjunto de vértices y E es el conjunto de enlaces. Asumimos que la red es bidimensional (es decir, cada nodo $v \in V$ esta embebido en el plano) y los nodos ad hoc se representan por los vértices del grafo. Cada nodo $v \in V$ tiene un radio de transmisión r. Sea $dist(v_1, v_2)$ la distancia entre los vértices $v_1, v_2 \in V$. Existirá un enlace entre dichos nodos $v_1, v_2 \in V$ iif $dist(v_1, v_2) \leq r$. Es decir, si y sólo si v_1 y v_2 pueden comunicarse directamente. En redes ad hoc algunos enlaces pueden ser unidireccionales debido al uso de diferentes radios de transmisión. Sin embargo, como el nivel de enlace puede detectarlos y ocultarlos al nivel de red, se suele asumir que los enlaces son bidireccionales. Es decir, $(v_1, v_2) \in E$ iif $(v_2, v_1) \in E$.

2.2. Formulación del problema

Dada una fuente multicast s y un conjunto de receptores R en una red representada por un grafo no dirigido, estamos interesados en encontrar el árbol multicast con un coste mínimo en cuanto al número de transmisiones que son necesarias para entregar un paquete desde s a todos los receptores $r_i \in R$. Para formular el problema apropiadamente necesitamos una serie de definiciones previas que damos a continuación.

Definición 1. Dado un grafo $G = (V, E)$, una fuente $s \in V$ y un conjunto de receptores $R \subset V$, definimos el conjunto T de todos los posibles árboles multicast en G que conectan la fuente s a todos los receptores $r_i \in R$. Podemos definir una función $C_t : T \rightarrow \mathbb{Z}^+$ tal que dado un árbol $t \in T$, $C_t(t)$ representa el número de transmisiones necesarias para entregar un mensaje desde la fuente a todos los receptores del árbol.

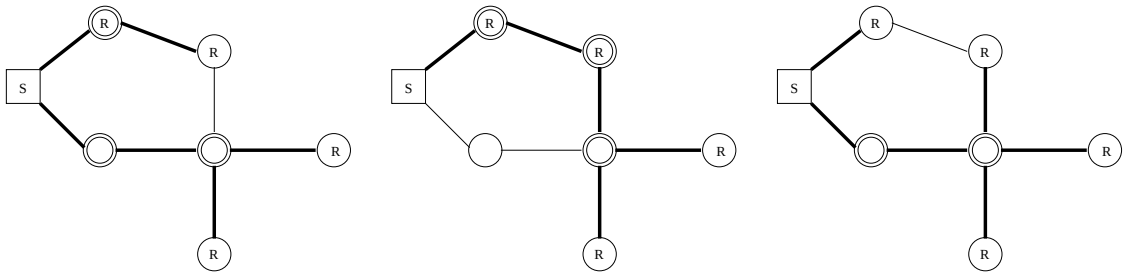
Lema 1. Dado un árbol $t \in T$ tal cual acabamos de definir, y si denotamos por F_t al conjunto de nodos que hacen de relay en t, entonces $C_t(t) = 1 + |F_t|$.

Demostración: Dado que en redes ad hoc el envío de mensajes se hace evitando duplicados, cada relay retransmitirá el mensaje enviado por s una única vez. Además, los nodos hoja no reenvían el mensaje. Por lo tanto, el número total de transmisiones necesarias es una de la fuente más una de cada uno de los relays, haciendo un total de $1 + |F_t|$. ■

Por lo tanto, tal cual vemos en el lema 1, para minimizar $C_t(t)$ el objetivo ha de ser reducir el número de nodos intermedios $|F_t|$.

Definición 2. Bajo las condiciones de la definición 1, y denotando por $t^* \in T$ al árbol multicast tal que $C_t(t^*) \leq C_t(t)$ para cualquier $t \in T, t \neq t^*$, definimos la sobrecarga de datos de un árbol $t \in T$ como $\omega_d(t) = C_t(t^*) - C_t(t)$. Obviamente, con esta definición $\omega_d(t^*) = 0$.

Basándonos en las definiciones anteriores, el problema se puede formular del siguiente modo. Dado un



(a) Shortest Path Tree, 4 Tx/packet (b) Steiner Tree, 4 Tx/packet (c) Min. Num. Tx, 3 Tx/packet

Figura 1: Diferencias en el coste de árboles multicast en redes ad hoc

grafo $G = (V, E)$, un nodo fuente $s \in V$, un conjunto de receptores $R \subset V$ y dado $V^{'} \subseteq V$ definido como $V^{'} = R \cup \{s\}$, encontrar un árbol $T^* \subset G$ tal que cumple las siguientes condiciones:

1. $T^* \supseteq V'$
2. $C_t(T^*)$ es mínimo

Al ser T^* un árbol, es obvio que es un grafo conexo, que combinado con la condición 1) establece que T^* ha de ser un árbol multicast. La condición 2) es equivalente a decir que su sobrecarga de datos ($\omega_d(T^*)$) es mínima y establece la optimalidad del árbol. Como mostramos en el siguiente teorema, este problema es NP-completo.

Teorema 1. Dado un grafo $G = (V, E)$, una fuente multicast $s \in V$ y un conjunto de receptores R, el problema de encontrar un árbol multicast $T^* \subseteq R \cup \{s\}$ tal que $C_t(T^*)$ es mínimo es NP-completo.

Demostración: De acuerdo con el lema 1, minimizar $C_t(T^*)$ es equivalente a minimizar el número de relays $F \subseteq T^*$. Por lo tanto, el problema es encontrar el conjunto más pequeño de relays F que conecta s con cada $r_i \in R$.

Si consideramos el caso particular en el que $R = V - \{s\}$, el objetivo sería encontrar el conjunto $F \subseteq T^*$ más pequeño que cubre al resto de nodos en el grafo ($V - \{s\}$). Este problema es famoso problema de la cobertura de vértices [12], que es NP-completo. Por lo tanto, al incluir un caso particular que es NP-completo, nuestro problema es también NP-completo. ■

En el siguiente teorema demostramos que en general el árbol con el menor costo (entendido como $C_e(t) = \sum_{e \in E} w(e)$) puede no ofrecer una menor sobrecarga de datos. Antes de presentar el teorema daremos algunas definiciones previas.

Definición 3. Bajo las mismas condiciones de la definición 1, y dado que cada enlace $e \in E$ tiene un costo asociado $w(e) > 0$, podemos definir la función $C_e : T \rightarrow \mathbb{Z}^+$, tal que dado un árbol $t \in T$, $C_e(t)$ es el costo del árbol definido como:

$$C_e(t) = \sum_{e \in E} w(e) \tag{1}$$

Para el caso particular de las redes ad hoc, podemos considerar que cada enlace tiene el mismo coste. De hecho, por simplicidad asumimos un coste unitario $w(e) = 1, \forall e \in E$. Incluso en este caso particular, el problema de calcular el árbol multicast T^* tal que $C_e(T^*)$ es mínimo (Steiner tree) sigue siendo NP-completo, tal cual demostró R. Karp en [7]. En este caso de costes unitarios, $C_e(T^*) = |V_{T^*}| - 1$ por definición de árbol, siendo $|V_{T^*}|$ el número de vértices del árbol T^*.

Teorema 2. Sea $G = (V, E)$ un grafo no dirigido, $s \in V$ una fuente multicast y $R \subseteq V$ el conjunto de receptores. El Steiner tree $T^* \subseteq G$ tal que $C_e(T^*)$ es mínimo podría no ofrecer una sobrecarga de datos mínima.

Demostración:

Para demostrar el teorema, mostraremos que dado un Steiner tree T^*, es posible encontrar un árbol T' tal que $C_e(T') \geq C_e(T^*)$ y $C_t(T') \leq C_t(T^*)$. Para ello, denotaremos por F^* y F' al número de relays en cada uno de los árboles. Al ser T' el árbol con la mínima sobrecarga de datos, se cumple (ver lema 1) la siguiente condición:

$$1 + F' \leq 1 + F^*$$

En un árbol multicast, el número de relays se puede dividir entre aquellos que son también receptores, y aquellos que no lo son. A éstos últimos se les suele llamar nodos "Steiner", y los denotaremos por $\mathbb{S}$. El número de relays que son también receptores será por tanto $|R - \mathbb{L}|$, siendo $\mathbb{L}$ el conjunto de nodos hoja (que por supuesto son receptores). Por lo tanto, la ecuación previa es equivalente a la siguiente:

$$1 + |\mathbb{S}'| + (|R| - |\mathbb{L}'|) \leq 1 + |\mathbb{S}^*| + (|R| - |\mathbb{L}^*|) \Rightarrow$$

$$|\mathbb{S}'| - |\mathbb{L}'| \leq |\mathbb{S}^*| - |\mathbb{L}^*| \Rightarrow$$
$$|\mathbb{S}'| - |\mathbb{S}^*| \leq |\mathbb{L}'| - |\mathbb{L}^*| \quad (2)$$

Además, por la propia definición de C_e, $C_e(T') \geq C_e(T^*) \Rightarrow |V'| - 1 \geq |V^*| - 1$. Por lo tanto, dado que el número de vértices es igual al emisor más el número de nodos Steiner más el número de receptores, podemos derivar la siguiente expresión:

$$|\mathbb{S}'| + |R| \geq |\mathbb{S}^*| + |R| \Rightarrow |\mathbb{S}'| \geq |\mathbb{S}^*|$$

Por lo tanto, de acuerdo con la ecuación 2, se puede construir un árbol T' tal que $C_t(T') \leq C_t(T^*)$. Para ello basta con conseguir tener más nodos hoja ($|\mathbb{L}'| - |\mathbb{L}^*|$), que el número de nodos Steiner ($|\mathbb{S}'| - |\mathbb{S}^*|$) adicionales que son necesarios. Un ejemplo de dicho árbol, que también prueba el teorema, se muestra en la Fig. 1. ■

3. Algoritmos propuestos

Dado que el problema que abordamos es NP-completo, en esta sección presentaremos dos algoritmos heurísticos para aproximar árboles multicast con una mínima sobrecarga de datos. Tal cual vimos en el teorema 2, una buena estrategia sería reducir el número de relays, a la vez que incrementar el número de receptores que quedan como nodos hoja. Las dos heurísticas que presentamos a continuación, se basan en esta idea.

3.1. Algoritmo greedy centralizado

Este primer algoritmo que presentamos, es útil para redes ad hoc inalámbricas en las que un nodo puede conocer la topología y calcular el árbol multicast.

Inspirado en los resultados del teorema 2, este algoritmo construye de forma sistemática sub-árboles de sobrecarga de datos mínima. Para ello, un nodo v es elegido como raíz del sub-arbol, si y sólo si cubre a dos o más nodos pendientes de ser conectados al árbol multicast. Dichos nodos pueden ser la fuente, los receptores o nodos raíz de otros sub-árboles ya construidos. El proceso termina cuando todos los sub-árboles quedan conectados. Este proceso, que mostramos en el algoritmo 1, comienza inicializando el conjunto de nodos a cubrir ("aux") a todos los receptores excepto aquellos directamente conectados a la fuente s. Inicialmente el conjunto de relays ("MF") se encuentra vacío. Tras la inicialización, el algoritmo repite el proceso de construir un sub-árbol, eligiendo el nodo v que cubre más nodos del conjunto "aux". Tras ello, v se añade al conjunto de relays "MF" y se añade también a "aux" al convertirse en un nodo a cubrir. Además, los receptores que cubre v, denotados por $Cov(v)$ se eliminan de la lista de nodos a cubrir. Este proceso se repite hasta que todos los nodos están cubiertos, o ya no sea posible elegir más nodos que conecten a 2 o más nodos del conjunto "aux". Si se da este último caso, entonces se calcula un Steiner tree entre las raíces de los diferentes sub-árboles. Es decir, entre los nodos que quedan en el conjunto "aux". Para hacer esto se puede usar cualquier heurística para Steiner trees, como por ejemplo el algoritmo MST que usamos nosotros en nuestras simulaciones.

Algoritmo 1 Mínima sobrecarga de datos centralizado.

```
1: MF ← ∅/ * relays * /
2: V ← V - {s}
3: aux ← R-Cov(s) + {s} / * nodosacubrir * /
4: repeat
5:    node ← argmax_{v∈V}(|Cov(v)|) s.t. Cov(v)≥2
6:    aux ← aux-Cov(v)+{v}
7:    V ← V-{v}
8:    MF ← MF + {v}
9: until aux = ∅ or node = null
10: if V!=∅ then
11:    Construir Steiner tree entre nodos en aux usando
       heurística MST
12: end if
```

Teorema 3. El algoritmo propuesto calcula un árbol multicast con una sobrecarga de datos menor o igual que la del Steiner tree.

Demostración: Consideremos el peor caso de nuestro algoritmo propuesto, en el que no se puede formar ningún sub-árbol. Existen dos posibles casos:

1. No hay receptores directamente conectados a la fuente. En ese caso Cov(s)=$\emptyset$ y el árbol resultante (T_1) es exáctamente el Steiner tree (T_2) que conecta a s y a todos los receptores. Por lo tanto, $C_t(T_1) = C_t(T_2)$.
2. Hay receptores directamente conectados a la fuente. En ese caso, el árbol resultante (T_1) es un Steiner tree entre la fuente s y todos los receptores excepto $Cov(s)$. Este árbol es un sub-árbol del Steiner tree (T_2) que conecta s con todos los receptores, por lo que $C_t(T_1) \leq C_t(T_2)$.

■

El algoritmo anterior es útil para algunos tipos de redes inalámbricas. Sin embargo, un algoritmo distribuido es mucho más recomendable para la gran mayoría de escenarios. En esta sección presentamos una versión distribuida del algoritmo anterior.

El algoritmo anterior consta de dos partes bien diferenciadas: (i) la construcción de los sub-árboles y (ii) la construcción de un Steiner tree entre las raíces de los sub-árboles.

Para construir el Steiner tree entre las raíces de los sub-árboles, en el esquema centralizado se suele usar la heurística MST. Sin embargo, al ser una heurística centralizada no nos sirve para el algortimo distribuido.

Dicha heurística comienza construyendo la "clausura métrica[1]" para el conjunto $\{s\} \cup R$ sobre todo el grafo, y después construye un MST para esa "clausura métrica". Finalmente cada enlace en el MST se sustituye por el camino más corto entre ambos nodos del grafo. La "clausura métrica" de un grafo es muy costosa de calcular de forma distribuida. Sin embargo, podemos aproximar dicha heurística MST de forma distribuida con el algoritmo 2, que mostramos a continuación.

La fuente (o la raíz del sub-árbol de la fuente) inunda un mensaje de solicitud de ruta (RREQ). Los nodos intermedios, al propagar el mensaje incrementan el campo "hop-count". Cuando el mensaje RREQ es recibido por una raíz de un sub-árbol, éste envía un mensaje de respuesta (RREP) hacia la raíz por el camino con menor número de saltos. Los nodos en ese camino se activarán como relays para ese grupo multicast. Además, un nodo raíz de un sub-árbol inicializará el contador de saltos a 0 al propagar el mensaje RREQ. Esto es precisamente lo que hace que nuestro algoritmo tenga un comportamiento similar a la construción del MST sobre la "clausura métrica". De hecho, se consigue el mismo efecto. Esto es, que cada raíz de un sub-árbol añadirá al árbol resultante el camino que le une a la raíz del subárbol donde se encuentra la fuente, o a la del sub-árbol más cercano (menor coste).

Algoritmo 2 Aproximación distribuida a MST

```
1: if thisnode.id = source − root then
2:    Enviar RREQ con RREQ.hopcount=0
3: end if
4: if recibido RREQ duplicado con mejor hopcount then
5:    prevhop ← RREQ.sender
6:    RREP.nexthop ← prevhop
7:    RREQ.sender ← thisnode.id
8:    if thisnode.isroot then
9:       envia(RREP)
10:      RREQ.hopcount ← 0
11:   else
12:      RREQ.hopcount++;
13:   end if
14:   send(RREQ)
15: end if
16: if recibido RREP y RREP.nexthop = thisnode.id then
17:   Activar MF_FLAG /* es_relay */
18:   RREP.nexthop ← prevhop
19:   envia(RREP)
20: end if
```

La segunda parte del algoritmo a hacer distribuida es la creación de los sub-árboles a conectar con el Steiner tree. Esta parte se puede realizar de forma local con sólo intercambiar unos pocos mensajes entre nodos vecinos. En concreto, al recibir el RREQ, los receptores inundan un mensaje Subtree_Join (ST_JOIN) sólo a sus vecinos, indicando el grupo multicast al que desean unirse. Estos vecinos responden con un Subtree_Join_Ack (ST_ACK) indicando el número de receptores que cubren. Esta información se conoce de forma local, contando el número de ST_JOIN recibidos en el paso anterior. Finalmente, los receptores envían también a sus vecinos un Subtree_Join_Activation (ST_JOIN_ACT) incluyendo el nodo que seleccionan como su padre, que es el vecino que cubre a un mayor número de nodos (≥ 2). Esta información también se sabe de forma local, tras comparar los ST_ACK recibidos. Aquellos nodos elegidos como padres se convierten en relays e inician de nuevo el proceso de enviar el ST_JOIN, pero esta vez los nodos que ya eligieron a un padre no responden a estos ST_JOIN.

En la siguiente sección mostramos como este algoritmo distribuido ofrece una aproximación casi tan buena como la del esquema centralizado, y ofrece un rendimiento mejor que el ofrecido por la heurística MST para calcular Steiner trees.

4. Simulación de resultados

Para evaluar la bondad de los algoritmos propuestos, los hemos simulado bajo diferentes condiciones. Los algoritmos considerados son las dos propuestas anteriores, así como la heurística MST para Steiner trees. Además, también simulamos árboles de camino más corto, que son los empleados por la gran mayoría de algoritmos para encaminamiento multicast en redes ad hoc actualmente.

4.1. Métricas de rendimiento

Nuestro objetivo es evaluar la optimalidad de la topología del árbol multicast producido por los diferentes algoritmos. Es por ello que usamos métricas diferentes a las comúnmente empleadas en redes ad hoc (como por ejemplo el ratio de entrega de mensajes, que depende enormemente de la tecnología inalámbrica considerada). En nuestro caso, las métricas a emplear son:

- Numero de transmisiones necesarias. El número total de mensajes enviados por la fuente o los relays para entregar un datagrama de la fuente a todos los receptores. Esta métrica sirve para evaluar la optimalidad del árbol como conjunto.
- Número medio de saltos. La media del número de saltos entre cada receptor y la fuente. Esta medida da una idea de la optimalidad individual de cada caminio.

Por lo tanto, considerando estas métricas junto a un nivel MAC perfecto (sin colisiones, retransmisiones o

[1]Subgrafo formado un subconjunto de nodos, conectados por enlaces cuyo coste es la suma del los costes de su camino más corto en el grafo original

interferencias), garantizamos una comparación equánime.

4.2. Metodología de simulación

Todos los enfoques han sido evaluados considerando un número diferente de receptores y una densidad variable de la red. En concreto, se ha considerado un total de 500 nodos, con un radio de transmisión de 250m. El número de receptores se ha variado entre un 1 y un 40 % del número total de nodos, que corresponde a un rango entre los 5 y los 200 receptores. La densidad de la red es variada entre 100 y 500 $nodos/Km^2$.

Para cada combinación de parámetros de simulación se han realizado un total de 91 ejecuciones con diferentes semillas aleatorias y grafos generados aleatoriamente, haciendo un total de más de 100000 simulaciones. El error mostrado en las figuras se ha obtenido usando un intervalo de confianza del 95 %.

4.3. Evaluación del rendimiento

En todas las figuras que mostramos a continuación, SPT se refiere al árbol de camino más corto, MST al Steiner tree basado en la heurística MST y por último MNT y MNT2 se refiere a los algoritmos propuestos, siendo MNT el enfoque centralizado y MNT2 el enfoque distribuido.

En la Fig. 2 mostramos para una red con densidad intermedia de 222 $nodos/Km^2$ (equivalente a un área de 1500x1500m) como varía el número de transmisiones necesarias respecto al número de receptores. Como era de esperar, cuando el número de receptores es menor que 20, todos los esquemas presentan un rendimiento similar. Esto se debe a que los nodos tienden a estar muy dispersos, haciendo menos probable la formación de sub-árboles. Sin embargo, tan pronto aumenta el número de receptores, la creación de dichos sub-árboles permite que los enfoques propuestos consigan reducir el número de transmisiones requeridas. Además, dado que el enfoque SPT no intenta minimizar el coste del árbol multicast, se aprecia como su rendimiento en este aspecto es bastante peor que el resto de alternativas. Respecto a los enfoques propuestos, se aprecia que el esquema distribuido (MNT2) ofrece un rendimiento un poco inferior al centralizado (MNT). Esto se debe a que a la hora de conectar las raíces de los sub-árboles emplea una aproximación a la "clausura métrica" mientras que el enfoque centralizado puede calcularla de forma totalmente precisa. Sin embargo, como se aprecia la diferencia no es muy grande. Además, se aprecia que ambos enfoques son capaces de ofrecer mejores resultados que los Steiner trees (MST).

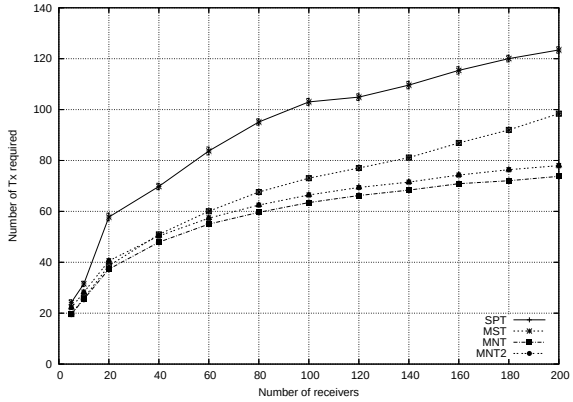


Figura 2: Número de transmisiones vs. Número de receptores

La Fig. 3 permite apreciar el impacto en la longitud de los caminos de los diferentes enfoques. Como era de esperar, el algoritmo SPT ofrece el mejor resultado. Esto se debe al hecho de que los otros algoritmos sacrifican la longitud de los caminos, para poder reducir el coste del árbol multicast. Como se puede apreciar esta métrica es mucho más variable al número de receptores que lo era el número de transmisiones necesarias. Es por ello que aparecen unas barras de error más grandes para MST, MNT y MNT2.

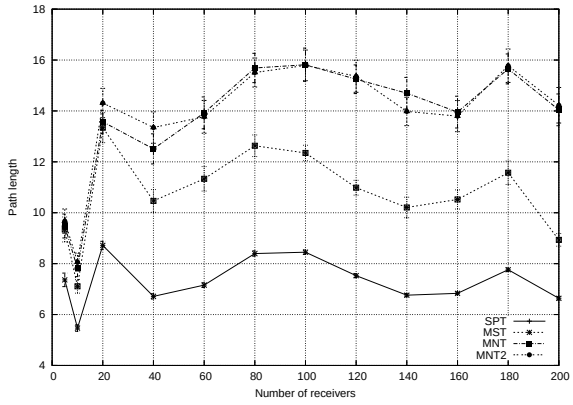


Figura 3: Longitud media de los caminos vs. Número de receptores

Otro aspecto importante a considerar es la variación del rendimiento respecto a la densidad de la red. Estos resultados son muy importantes para determinar en qué escenarios se comportan mejor los diferentes esquemas. En concreto, hemos considerado dos casos diferentes: un número medio de receptores (20 % de los nodos) y un alto número de receptores (36 % de los nodos). Como hemos mostrado antes, el caso de un número limitado de receptores no es interesante ya que todos los esquemas ofrecen un rendimiento similar.

En las Figs. 4 y 5 mostramos dichos resultados. Tal cual se muestra, a mayor densidad, mejor es el rendimiento de todos los enfoques. Esto se debe a

que al aumentar la densidad disminuye la longitud de los caminos, por lo que en general hace falta un número más reducido de transmisiones independientemente del algoritmo. Sin embargo, si comparamos los diferentes esquemas vemos que la reducción en el número de transmisiones es mayor en nuestros esquemas propuestos. Además, esa diferencia se hace mayor al aumentar la densidad de la red.

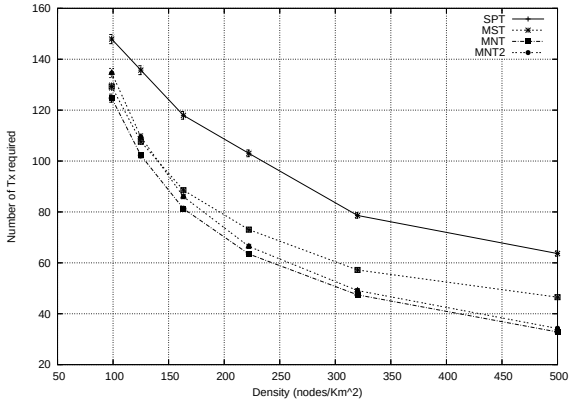


Figura 4: Número de TXs para 100 receptores vs. densidad de la red

Esto se explica por el hecho de que para mayores densidades, es más probable que varios receptores puedan estar cerca de un mismo nodo, facilitando por tanto la creación de sub-árboles de costo reducido.

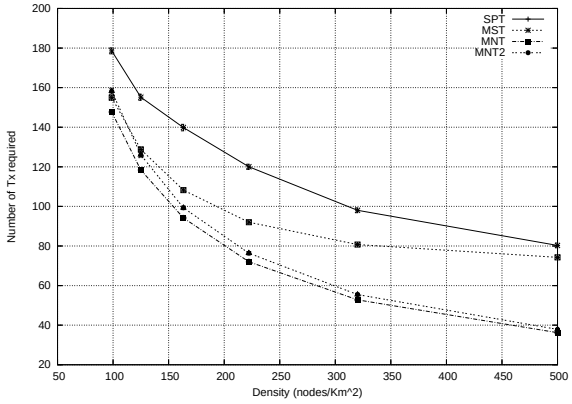


Figura 5: Número de TXs para 180 receptores vs. densidad de la red

Además, al aumentar la densidad, los rendimientos del enfoque centralizado y distribuido se acercan cada vez más. Esto se debe a que en redes densas, el número de saltos entre cualquier par de nodos también se reduce. Esto hace que la diferencia entre la "clausura métrica" y su aproximación en número de saltos se reduzca también. Es por ello que nuestro enfoque distribuido resulta muy interesante para redes densas tales como las redes de sensores, en las que el grado medio de un nodo suele ser muy alto.

Si comparamos las dos figuras podemos ver que la diferencia en el número de receptores sólo varía un poco las diferencias concretas de rendimiento entre enfoques. Sin embargo, la densidad de la red es la que tiene un mayor efecto en el rendimiento general de los algoritmos.

5. Conclusiones

En este artículo hemos mostrado que el Steiner tree, generalmente considerado como el árbol multicast de menor consumo de ancho de banda, no representa una solución optima en árboles multicast sobre redes inalámbricas ad hoc. Ello se debe a que la formulación original del Steiner tree no considera las reducciones de ancho de banda que pueden conseguirse en medios broadcast. Dadas esas limitaciones, hemos reconsiderado el problema de obtener el árbol multicast mínimo en dichas redes, en base a la optimización del número de transmisiones necesarias para enviar un datagrama desde una fuente multicast a todos los receptores.

Hemos demostrado que esta formulación es adecuada para redes ad hoc, y que este problema es NP-completo. Es por ello, que hemos propuesto nuevos algoritmos heurísticos para aproximar árboles multicast óptimos en este tipo de redes. Nuestras simulaciones demuestran que nuestros algoritmos ofrecen un mejor resultado que las aproximaciones para Steiner trees sobre una variedad de escenarios en terminos de la densidad de los receptores y de la red.

Los resultados muestran que a mayor densidad de la red, mejor rendimiento ofrecen nestros algoritmos en comparación con el resto de enfoques. Estos resultados son muy prometedores como una posible vía futura para tratar problemas similares en redes de sensores, que se caracterizan por densidades de red muy elevadas, y donde los árboles multicast inversos son comunes como topologías de recogida de datos.

Agradecimientos

Parte de este trabajo ha sido financiado por el MEC por medio del programa Ramón y Cajal, y el proyecto SAM (MCYT, TIC2002-04531-C04-03).

Referencias

[1] S. Deering, "Multicast Routing in a Datagram Internetwork," *Ph.D. Thesis, Electrical Engineering Dept., Stanford University*, Dec. 1991.

[2] S.-E. Deering and D.-R. Cheriton, "Multicast Routing in datagram internetworks and extended LANs," *Transactions on Computer Systems*, vol.8, no.2, May 1990, pp. 85–110.

[3] J. Moy, "Multicast routing extensions for OSPF," *Computer communications of the ACM*, vol.37, no.8, August 1994, pp.61–66.

[4] T. Ballardie, P. Francis and J. Crowcroft, "Core Based Trees (CBT) – An architecture for scalable inter-domain multicast routing, " *Proc. of ACM SIGCOMM'93*, San Francisco, CA, October 1993, pp.85–95.

[5] S. Deering, D.-L. Estrin, D. Farinacci, V. Jacobson, C.-G. Liu and L. Wei, "The PIM architecture for wide-area multicast routing," *IEEE/ACM Transactions on Networking*, vol.4, no.2, April 1996, pp. 153–162.

[6] C. Cordeiro, H. Gossain and D. Agrawal, "Multicast over Wireless Mobile Ad Hoc Networks: Present and Future Directions" *IEEE Network*, no. 1, Jan 2003, pp. 52–59.

[7] R.-M. Karp, "Reducibility among combinatorial problems," *In Complexity of computer computations*, Plenum Press, New York, 1975, pp.85–103.

[8] B.-M. Waxman, "Routing of Multipoint Connections," *IEEE Journal on Selected Areas in Communications*, vol. 6, no. 9, December 1998, pp. 1617–1622.

[9] L. Kou, G. Markowsky, and L. Berman, "A fast algorithm for Steiner trees," *Acta Informatica*, no. 15, vol. 2, 1981, pp. 141–145.

[10] J. Plesnik, "The complexity of designing a network with minimum diameter," *Networks*, no. 11, 1981, pp. 77–85.

[11] A. Zelikovsky, "An 11/6-approximation algorithm for the network Steiner problem," *Algorithmica*, no. 9, 1993, pp.463–470.

[12] S. Even, "Graph Algorithms," *Computer Science Press*, 1979, pp. 204–209.

[13] H. Lim and C. Kim, "Multicast Tree Construction and Flooding in Wireless Ad Hoc Networks," *Proceedings of the 3rd ACM international workshop on Modeling, analysis and simulation of wireless and mobile systems*, Boston, MA, USA. August, 2000, pp. 61–68.

Definición del comportamiento de gestión de red con reglas SWRL en un marco de gestión basado en ontologías en OWL

Antonio Guerrero[1], Víctor A. Villagrá[1], Jorge E. López de Vergara[2].
[1]Dpto. de Ingeniería de Sistemas Telemáticos, Universidad Politécnica de Madrid (DIT-UPM).
[2]Dpto. de Ingeniería Informática, Universidad Autónoma de Madrid.
antonio.guerrerocasteleiro@telefonica.es, villagra@dit.upm.es, jorge.lopez_vergara@uam.es

Abstract*. The goal of the ontology-based management is the improvement of current network management interoperability techniques through the study and application of formal ontologies. Prior research work has analyzed different aspects of the problem: a study of the semantic expressiveness of current management information definition languages, the application of ontology languages (DAML+OIL and OWL) to represent the management information definitions, and the mapping and merging processes to obtain a semantic integration of those definitions. The usage of an ontology language allows additional advantages: these ontologies can include the definition of behavior rules in order to reason with them. Thus, usual behavior definitions included implicitly in the management information definitions and explicitly in policy definitions can then be expressed formally, and included with those definitions, which allows a manager reasoning and working with them. This paper focuses in the definition of behavior rules in management information with SWRL, a rule language defined to complement OWL functionality.*

1 Introducción

La heterogeneidad de los recursos actuales en las redes y servicios de telecomunicaciones ha hecho que se definan diversos modelos de gestión integrada, los cuales permiten acometer la gestión de un entorno heterogéneo utilizando un mismo mecanismo de comunicaciones y acceso a la información de gestión. Inicialmente, surgieron dos modelos (modelo de gestión Internet y modelo de gestión OSI) y, aunque el modelo de gestión Internet ocupó un gran ámbito de actuación, el modelo de gestión OSI todavía es utilizado en algunos escenarios sobre todo orientados a la gestión de redes de telecomunicación con la arquitectura TMN.

Adicionalmente, a finales de la década de los 90, surge un nuevo modelo que utiliza los protocolos de la Web para acometer los mismos objetivos de gestión integrada (modelo WBEM – *Web Based Enterprise Management*, Gestión de Empresa Basado en Web), si bien este modelo no excluía a los otros dos sino que permitía una cierta integración de distintos modelos de gestión de red integrada.

En este escenario de utilización de múltiples modelos de gestión de red integrada surge la misma problemática que se intentó solucionar con dichos modelos (la gestión propietaria), ya que un gestor tiene que interactuar con distintos recursos utilizando distintos mecanismos. Pero incluso así, el gestor no puede realizar una gestión verdaderamente integrada de su entorno, ya que no puede discernir acerca de la semántica de recursos definidos en distintos modelos para poder aplicar una política de gestión común a sus recursos, de forma independiente al modelo bajo el cual estén definidos.

Para solventar esta cuestión, en [1] se realiza una propuesta de la denominada Gestión Semántica basada en Ontologías, que permite a un gestor trabajar con un único modelo de información que, teniendo en cuenta los aspectos semánticos, fusiona las distintas definiciones de recursos gestionados realizados en modelos diferentes. En dicha propuesta, se menciona otra ventaja que tiene esta aproximación, que es permitir integrar dentro de dicho modelo unificado la definición de aspectos de comportamiento del gestor y de los recursos, los cuales suelen estar mencionados como comentarios en las definiciones o explícitamente declarados pero fuera de las definiciones de los recursos gestionados.

Este artículo propone un punto de partido para dicha integración: comenzando con un modelo de información unificado y definido en un lenguaje de ontologías como es OWL, se propone la inclusión en dicho modelo de definiciones de comportamiento utilizando el Lenguaje de Reglas de la Web Semántica (SWRL, *Semantic Web Rule Language*). Para ello, el siguiente apartado presenta la arquitectura de gestión semántica sobre la que se va a trabajar, presentando a continuación el lenguaje SWRL. Dicho lenguaje permitirá la definición de comportamiento dentro de la información de gestión definida, como se comenta más adelante. Al mismo tiempo, se explican los tipos de comportamiento analizados: restricciones implícitas, comportamiento explícito del gestor y comportamiento explícito de los elementos gestionados. Para cada uno de estos tipos se aportan ejemplos de la aplicación de SWRL a la definición de comportamiento. Finalmente se resumen las ideas más relevantes.

2 Gestión Semántica

La arquitectura global propuesta en [2] se basa en un gestor que trabaja y razona con un único modelo de información de gestión representado mediante ontologías. Este gestor maneja elementos de diferentes dominios (SNMP, CIM, etc.) desde un punto de vista común y neutral a todos ellos. La Fig. 1 muestra la arquitectura propuesta para el gestor semántico.

El objetivo principal es que el gestor utilice un único modelo de información, pero en muchos casos los distintos recursos de red están definidos en distintos dominios de gestión (MIBs de SNMP, esquemas CIM, etc.) y hay que acceder a ellos utilizando los protocolos definidos para dicho dominio. Por lo tanto es necesario traducir e integrar estas definiciones en una definición unificada, teniendo en cuenta la semántica de dichas definiciones. Es decir, un mismo recurso, definido dos veces en dos modelos diferentes, debe quedar con una única representación en el modelo unificado, y con dos traducciones a los modelos origen. Se trata por tanto no sólo de una tarea de traducir definiciones sintácticamente, sino también de integrarlas desde un punto de vista semántico.

Las ontologías de correspondencia resultantes son utilizadas por los llamados "proveedores" o pasarelas de la Fig. 1 para traducir la información del gestor en el modelo unificado, a la información de los elementos de red en sus particulares modelos de gestión.

La utilización de un lenguaje de ontologías para definir la información de gestión ofrece ventajas adicionales, la primera de las cuales es la posibilidad de utilizar herramientas existentes para trabajar y razonar con las ontologías (por ejemplo, motores de inferencia utilizados en inteligencia artificial).

Otra ventaja, comentada en el apartado anterior, es que la ontología de gestión puede llegar a incluir la definición de reglas de comportamiento de la información de gestión. De esta forma, las definiciones de comportamiento que se suelen incluir implícitamente en las definiciones de la información de gestión (se indican en lenguaje natural, o se suponen), pueden llegar a ser expresadas de forma integrada con las definiciones de información de gestión, y en su mismo lenguaje (lenguaje de ontologías). Es decir: esta aproximación supone que las definiciones de comportamiento quedan ahora expresadas formalmente en el mismo lenguaje de información de gestión, y pueden ser interpretadas y validadas por un gestor semántico (basado en ontologías) para trabajar y razonar sobre ellas. Todas las definiciones de gestión, tanto las de la información de gestión (MIBs) como las reglas de comportamiento, se integran ahora en una misma ontología de gestión de red.

Como lenguaje de definición de ontologías, se propone el uso de OWL [3], el cual es un lenguaje de ontologías de propósito general definido para la Web Semántica que contiene todas las construcciones necesarias para describir formalmente la mayor parte de las definiciones de información de gestión: clases y propiedades, con jerarquías, y restricciones de rango y de dominio [4]. SWRL [5] extiende el conjunto de axiomas de OWL para incluir reglas condicionales (cláusulas de Horn), de tipo *si…entonces...*

Los axiomas y las reglas pueden ser utilizados en este marco de gestión para:

1) Restringir o definir de forma aún más precisa el comportamiento de la información de gestión en OWL. Esto permite asegurar un correcto uso e implementación de la información de gestión.

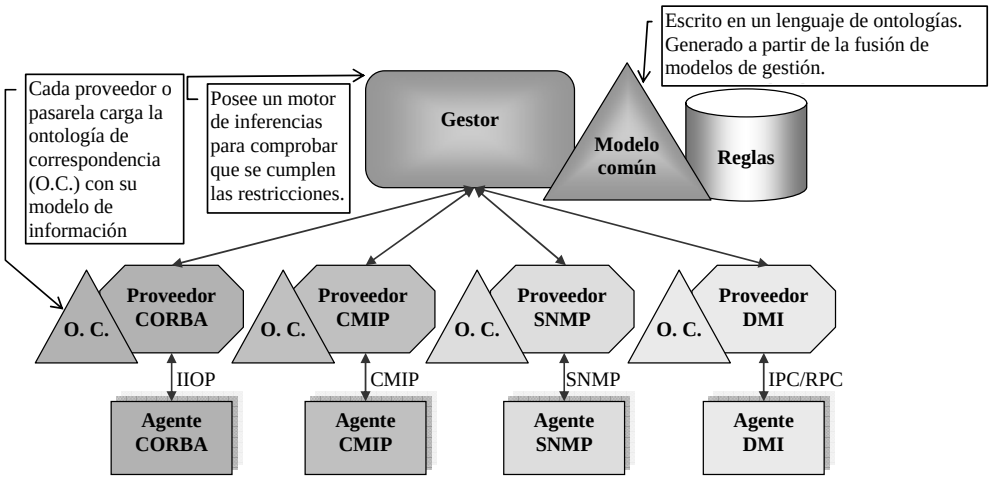


Fig.1: Arquitectura del gestor semántico

2) Definir formalmente el comportamiento del gestor. Esto permite indicar lo que debería hacer el gestor si se cumplen determinadas condiciones en los elementos gestionados.

3) Definir formalmente el comportamiento de los elementos gestionados. Esto permite indicar lo que deberían hacer los recursos gestionados ante eventos determinados.

El objetivo de este trabajo es integrar la definición de la información de gestión (expresada en OWL) con la definición del comportamiento de gestión (expresado en SWRL). Para ello, el siguiente apartado presenta SWRL como un lenguaje que permite describir reglas para ontologías descritas en OWL.

3 SWRL: Definición de reglas para ontologías den OWL

Una ontología en OWL contiene una secuencia de hechos y axiomas. Los axiomas pueden ser de varias clases, p.e., axiomas de subclase, axiomas de equivalencia de clases, restricciones sobre propiedades. SWRL propone extenderlos con axiomas de regla.

Un axioma de regla consiste en un *antecedente* (cuerpo) y un *consecuente* (cabeza), cada uno de los cuales está compuesto por un conjunto (puede ser vacío) de *átomos*.

En la sintaxis inteligible por seres humanos de SWRL, una regla tiene la siguiente forma:

```
antecedente ⇒ consecuente
```

De manera informal, una regla puede interpretarse como una indicación de que si el antecedente es cierto, entonces el consecuente también es cierto. Utilizando esta sintaxis (notación lógica clásica), la propiedad de "ser tío de" se escribiría de la siguiente forma:

```
Persona(?x) ∧ esPadre(?x,?y) ∧
esHermano(?y,?z) ⇒ esTio(?x,?z)
```

Una regla SRWL tiene por tanto la forma de una relación de implicación entre la *cabeza* y el *cuerpo*. La especificación de SWRL [5] ofrece una sintaxis abstracta que extiende la sintaxis abstracta de OWL descrita en [6] para incluir esta nueva relación en el lenguaje de ontologías. Las marcas XML que permiten describir estas reglas incluyen:

- `<ruleml:imp>`: Es el elemento que permite relacionar el *cuerpo* de la regla con la *cabeza*.

- `<ruleml:_body>`: Es el elemento que lista los *átomos* del *cuerpo* de la regla.

- `<ruleml:_head>`: Es el elemento que lista los *átomos* de la *cabeza* de la regla.

- `<ruleml:var>`: Permite definir las variables sobre las que evaluar las reglas.

- `<swrlx:individualPropertyAtom>`: Permite definir átomos referidos a propiedades concretas. También es posible definir átomos referidos a clases, rangos de datos, propiedades valuadas, o funciones propias de tipo matemático, cadenas de caracteres o fechas.

Como se ve, muchas de estas marcas no están definidas dentro del espacio de nombres de SWRL, sino de RuleML [7], un lenguaje de reglas definido con anterioridad que se ha tomado como base en la definición de SWRL. SWRL aporta sobre todo la definición de los átomos y su integración dentro de una ontología escrita en OWL.

4 Especificación del Comportamiento de Gestión en OWL+SWRL

4.1 Tipos de comportamiento de gestión

Para poder especificar comportamientos de gestión resulta conveniente clasificar qué tipos pueden existir. En concreto, se han identificado tres tipos de comportamiento de gestión:

1) Restricciones y reglas de comportamiento implícitas a los objetos modelados en las MIBs y esquemas CIM.

2) Comportamiento explícito del gestor, según la arquitectura tradicional gestor-agente en el que se indica como se comporta el gestor al obtener y analizar información de los agentes.

3) Políticas definidas explícitamente para especificar el comportamiento o configuración dinámica de los recursos gestionados (Gestión Basada en Políticas o *Policy-Based Management* [8])

En este marco de gestión semántica, las políticas se pueden definir en el mismo lenguaje de gestión, OWL+SWRL, que los objetos definidos en la base de información de gestión (MIB), con la ventaja que supone el trabajar con un modelo unificado.

Volviendo a la arquitectura de gestión propuesta en la Fig. 1, las definiciones de restricciones, reglas y políticas se almacenarían en la base de información de reglas (parte integrante del modelo común), cuyo

propósito es verificar la integridad de la información, y automatizar el control de los elementos gestionados.

Los siguientes apartados se enfocan en cada uno de los tipos de comportamiento identificados.

4.2 Restricciones implícitas sobre la información de gestión

Este tipo de reglas se refieren a restricciones en el tipo de datos, cardinalidad o acceso. Son restricciones típicas sobre las propiedades y clases de los objetos gestionados. En este caso, las reglas en SWRL permiten complementar a los mecanismos ya existentes en OWL para realizar las definiciones de información de gestión, de forma que permiten expresar restricciones más complejas: valores que dependen del valor de otras variables, relaciones entre objetos, comportamiento de máquina de estados de los valores, etc.

En general, SWRL permite representar formalmente restricciones de comportamiento que se expresen en lenguaje natural de forma condicional (si … entonces …). Esto incluiría, entre otros, dependencia de valores, comportamiento de estados, comportamiento temporal o tipos complejos como funciones compuestas. A continuación se presentan, algunos ejemplos de este tipo de comportamiento encontrados en MIBs SNMP y cómo se aplica SWRL para definir este tipo de restricciones. En muchos casos la definición de las restricciones requerirá de la creación de nuevas clases y propiedades que extenderían la ontología de gestión.

Ejemplo 1

El primer ejemplo muestra cómo especificar con SWRL que el valor de una variable dependa del valor de otra. Para ello, se hace uso de una definición recogida de la MIB II de SNMP, relativa a la máscara de una dirección de destino en una tabla de encaminamiento:

```
ipRouteMask OBJECT-TYPE
    SYNTAX  IpAddress
    ACCESS  read-write
    STATUS  mandatory
    DESCRIPTION
        "… If the value of the
ipRouteDest is 0.0.0.0 (a default
route), then the mask value is also
0.0.0.0.  ..."
    ::= { ipRouteEntry 11 }
```

Esta restricción implícita se expresa en lenguaje natural en la cláusula `DESCRIPTION`, pero no se define formalmente en SMIv2, por lo que un gestor no puede implementarla de manera automática al compilar la MIB.

Suponiendo que la MIB II hubiese sido traducida e integrada en la base de información de gestión en OWL, la regla SWRL que definiría esta restricción sería la siguiente:

```
ipRouteEntry(IR?) ∧ swrlb:equal
(ipRouteDest(IR?), "0.0.0.0")

⇒ swrlb:equal (ipRouteMask(IR?),
"0.0.0.0")
```

donde *ipRouteDest* e *ipRouteMask* son propiedades de la clase *ipRouteEntry*.

Cabe señalar que esta restricción aquí definida de forma simple, no es posible expresarla en otros lenguajes de definición de información de gestión. De hecho, el SMIng Working Group del IETF propuso el siguiente objetivo para la siguiente generación del lenguaje SMI, incluido en [9]: "*SMIng should provide mechanisms to formally specify constraints between values of multiple attributes*", pero su desarrollo fue desestimado: "*This objective as is has been rejected as too general, and therefore virtually impossible to implement*".

Por otro lado, y reforzando la idea en estudio, ocurre que tampoco es posible expresar esta restricción en OWL sin SWRL: una restricción sobre una propiedad (cláusula *owl:restriction*) no puede hacerse depender del valor de otra propiedad del mismo objeto.

Ejemplo 2

Este segundo ejemplo, obtenido también de la MIB II de SNMP, puede servir para mostrar la definición del comportamiento de estados en SWRL. En este caso se toma la columna de la tabla de conexiones de TCP que indica el estado de cada conexión.

```
tcpConnState OBJECT-TYPE
    SYNTAX  INTEGER {
                closed(1),
                listen(2),
                synSent(3),
                synReceived(4),
                established(5),
                finWait1(6),
                finWait2(7),
                closeWait(8),
                lastAck(9),
                closing(10),
                timeWait(11),
                deleteTCB(12)
            }
    ACCESS  read-write
    STATUS  mandatory
    DESCRIPTION      "…"

    ::= { tcpConnEntry 1 }
```

Como este parámetro es de tipo read-write, se trata entonces de definir formalmente en SWRL la máquina de estados TCP que se ilustra en la Fig. 2 para forzar al gestor a cumplirla.

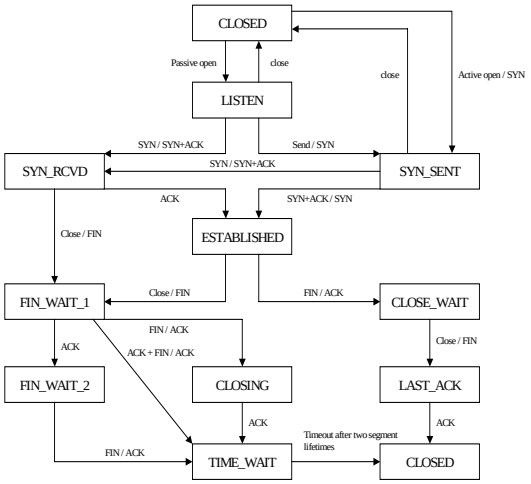


Fig.2: Máquina de estados TCP

En el modelo de información se ha definido una clase de objeto gestionado llamada *tcpConnEntry* con la propiedad relativa al estado actual de la conexión *tcpConnState*, y las propiedades auxiliares *tcpConnPreviousState*, *tcpConnNextState*, que indican la lista de estados previos y siguientes, respectivamente. Un ejemplo de una de las reglas que permitirían definir la máquina de estados es el siguiente:

```
tcpConnEntry(cx?) ∧
swrlb:equal(tcpConnState(cx?),
"fin_wait_1")

⇒ swrlb:member(tcpConnNextState(cx?),
nextStatesForFin_Wait_1_List)
```

donde *nextStatesForFin_Wait_1_List* sería una lista (*rdf:list*) con los valores "closing", "time_wait" y "fin_wait_2".

Ejemplo 3

El tercer ejemplo muestra un caso de restricción temporal, tomando para ello la columna de la tabla de interfaces de red que indica el último cambio de una interfaz. Este ejemplo de restricción temporal podría ser desarrollado en SWRL si se hace uso de otras clases y propiedades. A continuación se muestra la definición de la propiedad anteriormente comentada, donde se ha marcado en negrita la cláusula que podría representarse con SWRL.

```
ifLastChange OBJECT-TYPE
    SYNTAX  TimeTicks
    ACCESS  read-only
    STATUS  mandatory
    DESCRIPTION
        "The value of sysUpTime at the
time the interface entered its current
operational state.  If the current state
was entered prior to the last re-
initialization of the local network
management subsystem, then this object
contains a zero value."
    ::= { ifEntry 9 }
```

En este caso, tomando la propiedad *sysUpTime* de la clase *system*, se podría hacer dicha comparación temporal.

```
system(?x) ∧ ifEntry(?y) ⇒
swrlb:lessThan(ifLastChange(y?),
sysUpTime(?x))
```

Ejemplo 4

Un ejemplo de función compuesta a representar en SWRL puede ser un indicador sobre un determinado umbral, por ejemplo de temperatura (*flagTemp*):

```
Temperatura(t?) ∧  swrlb:lessThan(t?, 40) ⇒
flagTemp(ft?) ∧  swrlb:equal(valor(ft?), 0)

Temperatura(t?) ∧
swrlb:greaterThanOrEqual(t?, 40) ⇒
flagTemp(ft?) ∧  swrlb:equal(valor(ft?), t?)
```

Este indicador valdría 0º por debajo de 40º, y tomaría el valor de la temperatura en 40º o por encima.

4.3 Comportamiento explícito del gestor

El comportamiento del gestor ante el cumplimiento de determinadas condiciones en la red o los sistemas, puede especificarse por medio de reglas condicionales del tipo **if** *condición* **then** *acción*, que en SWRL se convierten en

condición ⇒ *acción*

o, ampliando la definición

conjunto de condiciones ⇒ *conjunto de acciones*

Acciones

Una de las carencias que posee OWL es la definición de funciones o métodos, habituales en lenguajes de información de gestión tales como CIM o GDMO. Para solventar este problema, las acciones en OWL pueden ser representadas mediante la ontología de servicios OWL-S (*OWL-Services*) [10]. Dicha ontología define entre otras cuestiones la clase *Process* que puede ser utilizada para este fin. Si se hace el símil de que ejecutar una acción puede realizarse "llamando a un servicio" que ejecuta la acción, entonces se puede definir esa acción como un *proceso OWL-S*, y crear una instancia de la clase *perform* con dicho *proceso* como argumento:

Perform(MyProcess)

La clase *Perform* es una clase auxiliar definida para representar la ejecución de procesos atómicos dentro de un proceso compuesto.

Ejemplos: `Perform(Reset)`, `Perform(SendAlarm(…))`, `Perform(SetIPRoute(…))`

Estos procesos pueden ser nuevas acciones que se definan, o bien ser el resultado la integración de las operaciones ya existentes en los modelos de información traducidos (procedimientos CIM, operaciones de tipo "set" de SNMP, etc.)

En la arquitectura gestor-agente propuesta en la Fig. 1, el gestor invocaría las llamadas a los servicios que ofrecerían los proveedores, y estos a su vez ejecutarían las operaciones correspondientes sobre los elementos gestionados. El siguiente ejemplo muestra este tipo de definición de comportamiento del gestor semántico ante determinadas condiciones en la red.

Ejemplo 5

Este ejemplo toma la clase *CIM_SystemDevice* del esquema CIM, que posee dos ejemplares asociados a dos puertos, y activa uno si el otro no funciona, siguiendo la regla:

If LogicalPort #1 is "Operatively Down", then enable LogicalPort #2

En SWRL:

```
CIM_SystemDevice(LP1?) ∧
swrlb:equal(deviceName(LP1?), "Lport1") ∧
CIM_SystemDevice(LP2?) ∧
swrlb:equal(deviceName(LP2?), "Lport2")  ∧
swrlb:equal(StatusInfo(LP1?),
"OPERATIVELY_DOWN")

⇒ Perform(SetAdminAvailability(LP2?,
"ENABLE"))
```

De esta forma se aplica una regla, no sobre toda una clase de objetos, sino sobre determinados objetos (ejemplares de una clase).

Al igual que ocurría con las restricciones implícitas, tampoco este tipo de definiciones explícitas podrían representarse en OWL sin SWRL como una restricción a los valores de una propiedad que dependan del valor de otra propiedad. A diferencia del ejemplo 1 del apartado anterior, en este caso se relacionan propiedades de objetos diferentes, que son ejemplares de una misma clase.

4.4 Comportamiento explícito de los elementos gestionados: aplicación a la gestión basada en políticas

Si las reglas que se definen en el modelo de información no son reglas para el comportamiento del gestor, sino reglas que definen el comportamiento de los elementos gestionados, entonces se necesita una arquitectura que permita distribuir las reglas o políticas a dichos elementos gestionados. Este tipo de arquitecturas son las que se definen en el entorno de PBM (*Policy-Based Management*) o gestión de red basada en políticas. En este trabajo se hará referencia a los siguientes elementos de la arquitectura PBM propuesta por el IETF/DMTF [7]:

- Dispositivos PEP (Policy Enforcement Point): son aquellos elementos que pueden aplicar o ejecutar las políticas.
- Dispositivos PDP (Policy Decision Point): actúan de intermediarios entre los PEPs y el repositorio de políticas, siendo los responsables de interpretar las políticas del repositorio y comunicar las acciones a los PEPs.
- Repositorio de políticas: almacena las políticas definidas, que serán distribuidas a los PDPs

Una arquitectura basada en políticas presenta las siguientes características clave:

- Centralización: la definición del comportamiento se realiza en un único punto y puede ser distribuida de forma masiva a la red en lugar de definirla y aplicarla para cada elemento de forma separada
- Niveles de abstracción: se pueden definir políticas a diferentes niveles: alto nivel (p.e. reglas de negocio), niveles intermedios (p.e. definiciones en el nivel de servicio), y bajo nivel (p.e. políticas que aplican los elementos de red). Se prevé por tanto la necesidad de traducciones entre niveles para convertir las reglas de alto nivel en las políticas de bajo nivel que aplicarán los elementos de red, para lo cual resulta útil disponer de modelos (modelos de negocio, modelos de servicio, modelos de red). Los lenguajes de definición de políticas existentes pueden estar más orientados hacia la definición de políticas en unos u otros niveles. Como ejemplos citaremos PONDER y RBAC que son lenguajes que llamaríamos de "alto nivel" porque el modelo de información es muy cercano a la forma de pensar del ser humano, mientras que PCIM y PIB de COPS-PR permiten expresar políticas de muy bajo nivel, más cercanas a los lenguajes de gestión de los elementos de red.

A continuación se realiza una primera aproximación a la posibilidad de utilizar ontologías en

OWL+SWRL como lenguaje para la definición de estos tipos de políticas, mediante los ejemplos siguientes.

Ejemplo 6: política de alto nivel en PONDER

En este ejemplo se define en SWRL una política ya definida en el lenguaje de definición de políticas PONDER, extraída de [11].

```
type rel ReportingT (ProjectManagerT pm,
SecretaryT secr) {
        inst oblig reportWeekly {
           on timer.day ("monday") ;
           subject secr ;
           target pm ;
           do mailReport() ;
        }
        // . . . other policies
}
```

Esta política obligatoria llamada *reportWeekly* especifica que el sujeto con rol *SecretaryT* debe enviar por correo un informe cada lunes al sujeto con rol *ProjectManager*. En SWRL se podría expresar de la siguiente manera:

```
ProjectManagerT(pm?) ∧ SecretaryT(secr?) ∧
Timerday(t?) ∧ swrlb:equal (t?, "monday")

⇒ Perform(mailReport(secr?, pm?))
```

De esta forma, otros tipos de políticas de PONDER u otros lenguajes de definición de políticas como PCIM y COPS podrían ser definidas en SWRL: políticas de autorización, de obligación, de filtrado, etc.

Ejemplo 7: política de bajo nivel

A continuación se presenta una definición en SWRL de una política de bajo nivel que podría ser ejecutada directamente por un elemento de red.

La regla siguiente marcaría con "0100" el campo "ToS" (tipo de servicio) de cada paquete IP cuya dirección IP destino cumpliese la condición de pertenecer al rango indicado.

```
IPpacket(ippacket?) ∧
InsideIPRange(DestAddress(ippacket?),
"192.168.1.0 – 192.168.1.255")

⇒ Perform(SetTOSlabel(ippacket?, "0100"))
```

Este ejemplo es tal vez excesivamente sencillo y de escasa utilidad, puesto que posiblemente realizando una correspondencia directa de la PIB (*Policy Information Base*), en caso de existir, o del lenguaje de definición de políticas de bajo nivel que maneje el propio elemento de red o PDP fuera la mejor forma de representar las políticas de bajo nivel en la ontología común de información de gestión. Para esto se abordaría la traducción e integración en la ontología común, por un lado de la lista de condiciones que disparan la política (<*condition set*>), y por otro de la lista de acciones (<*action list*>) que ejecuta la política. Para ello puede resultar útil la aplicación del mecanismo M&M (*Merge and Map*, Fusión y Correspondencia) [2] de traducción e integración de información de diferentes lenguajes de gestión a la ontología común. El resultado de la aplicación de este método permitiría que, a la inversa, pudiesen traducirse las políticas representadas en la ontología en OWL+SWRL al lenguaje nativo que entiendan los PDPs.

Nótese que las políticas mostradas son políticas sin eventos del tipo "<*condition set*> *then do* <*action list*>". No existen eventos explícitos para la evaluación de las condiciones, sino que los agentes deben realizar esta tarea cuando existe algún evento implícito como, por ejemplo, el inicio de una sesión. Este sería el caso por ejemplo de PCIM, pero no de COPS y PONDER que sí permiten políticas del tipo *event-condition-action*.

La gestión basada en políticas también trata sobre la traducción de políticas entre diferentes niveles de abstracción, de forma que las políticas definidas en los niveles más altos (p.e. nivel de negocio y nivel de servicio) son traducidas a políticas de más bajo nivel (p.e. nivel de red y de sistemas). El hecho de que las definiciones de políticas de los distintos niveles en una misma ontología de gestión estén integradas en un mismo lenguaje (OWL+SWRL) puede facilitar la traducción de políticas entre niveles, aunque este estudio queda fuera del alcance del presente trabajo.

5 Conclusiones

Como se ha mostrado en estudios previos [1][2][4] los lenguajes de ontologías tales como OWL incluyen las construcciones necesarias para definir las características típicas de la información de gestión de red que pueden encontrarse en otros lenguajes de definición de gestión como SMI-SNMP, MOF-CIM, etc., y que permite traducir e integrar las definiciones de los distintos lenguajes desde una punto de vista semántico.

En este marco de gestión semántica, este trabajo ha mostrado como SWRL:

1) añade capacidad de expresión para definir restricciones y reglas para las definiciones de la información de gestión expresadas en la ontología común en OWL propuesta. SWRL permite expresar formalmente restricciones y reglas de definición del comportamiento más complejas que las que podían definirse en OWL, enriqueciendo y

añadiendo mayor expresividad a las definiciones de gestión.

2) permite definir explícitamente el comportamiento del gestor, y de los objetos gestionados, en el mismo lenguaje de ontologías, OWL+SWRL, que se ha propuesto para las definiciones de la información de gestión. Esto incluye:

   - Acciones que ejecutará el gestor ante la detección de determinadas condiciones en los elementos gestionados o en el propio gestor
   - Acciones que ejecutarán los elementos de red o sistemas gestionados ante la existencia de determinadas condiciones (políticas que definen el comportamiento de los elementos gestionados)

No obstante quedaría aún pendiente la integración en este marco de gestión semántico, de otros mecanismos que presentan los lenguajes de definición de políticas tales como eventos de red, eventos temporales, identificación y clasificación de políticas, etc.

Por otro lado, existe el problema de definir formalmente el conjunto de restricciones que actualmente hay definidas de manera implícita o explícita. En este sentido puede ayudar el uso de herramientas que integren la recogida, fusión y correspondencia de información de gestión con la búsqueda mediante heurísticos de restricciones descritas en los campos de descripción (con alta probabilidad cadenas como “if*then”, “have to” o “must” formarán parte de una de estas restricciones), y su posterior definición en SWRL a partir de la información ya definida en OWL.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el Ministerio de Educación y Ciencia a través del proyecto GESEMAN (TIC2002-00934).

Referencias

[1] Jorge E. López de Vergara, Víctor A. Villagrá, Juan I. Asensio, Julio Berrocal, “Ontologies: Giving Semantics to Network Management Models”, IEEE Network, Vol. 17, No. 3, May/June 2003. ISSN 0890-8044.

[2] Jorge E. López de Vergara, Víctor A. Villagrá, Julio Berrocal, “Benefits of Using Ontologies in the Management of High Speed Networks”, Lecture Notes in Computer Science, Vol. 3079, Springer-Verlag. ISSN 0302-9743

[3] Michael K. Smith, Chris Welty, Deborah L. McGuinness, “OWL Web Ontology Language Guide”, W3C Recommendation 10 Feb. 2004.

[4] Jorge E. López de Vergara, Víctor A. Villagrá, Julio Berrocal, “Applying the Web Ontology Language to management information definitions”, IEEE Communications Magazine, Vol. 42, Issue 7, July 2004, pp. 68-74. ISSN 0163-6804

[5] Ian Horrocks, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosof, Mike Dean, “SWRL: A Semantic Web Rule Language Combining OWL and RuleML”, W3C Member Submission 21 May 2004.

[6] Peter F. Patel-Schneider, Patrick Hayes, Ian Horrocks, “OWL Web Ontology Language Semantics and Abstract Syntax”, W3C Recommendation 10 February 2004.

[7] Rule Markup Initiative, http://www.ruleml.org/

[8] A. Westerinen, J. Schnizlein, J. Strassner, M. Scherling, B. Quinn, S. Herzog, A. Huynh, M. Carlson, J. Perry, S. Waldbusser, “Terminology for Policy-Based Management”, IETF Request For Comments 3198

[9] C. Elliott, D. Harrington, J. Jason, J. Schoenwaelder, F. Strauss, W. Weiss, “SMIng Objectives”, IETF Request For Comments 3216, December 2001.

[10] David Martin, editor, “OWL-S: Semantic Markup for Web Services”, W3C Member Submission 22 November 2004

[11] Nicodemos Damianou, Naranker Dulay, Emil Lupu, Morris Sloman, “The PONDER Policy Specification Language”, Proc. Policy 2001: Workshop on Policies for Distributed Systems and Networks, Bristol, UK, 29-31 Jan. 2001, Springer-Verlag LNCS 1995, pp. 18-39.

AVATAR: Un sistema de recomendación personalizada de contenidos televisivos basado en información semántica*

Manuel Ramos Cabrer, Yolanda Blanco Fernández, Alberto Gil Solla
Departamento de Ingeniería Telemática. Universidad de Vigo
ETSI de Telecomunicación. C/ Maxwell s/n. Campus Universitario.
36310 - Vigo (Pontevedra)
Teléfono: 986 81 38 70 Fax: 986 81 21 16
E-mail: mramos,yolanda,agil@det.uvigo.es

Abstract *In this paper we present a recommender system of personalized TV contents, named AVATAR, for which we propose a muti-agent architecture, that combines different knowledge inference strategies. We focus on the description of one of these strategies, the semantic one, explaining an example in the context of personalized digital television. Our goal is to take advantage of the experience of the Internet, from the current syntactic search engines to the newest results in Semantic Web area. Trying to be always under the umbrella of the most used standards, we adopt the TV-Anytime specification for metadata definition of TV contents, the OWL language to build a suitable ontology for knowledge representation, and the MHP specification for system development.*

1. Introducción

En la actualidad estamos asistiendo a un cambio fundamental en la TV: la migración de la TV analógica a la digital. Este cambio tiene dos implicaciones principales: un considerable aumento en la capacidad para difundir más canales en el mismo ancho de banda, y la posibilidad de enviar aplicaciones software conjuntamente con los contenidos audiovisuales, posibilitando la aparición de nuevas oportunidades de negocio [7].

En este nuevo escenario, los usuarios podrán acceder a un mayor número de canales de distintos proveedores y a muchos más contenidos que hoy en día. Por ello, serán necesarias nuevas herramientas que ayuden a los usuarios en la selección de contenidos que pueden ser de su interés. En el campo de la TV, los sistemas recomendadores pueden permitir al usuario jugar un papel activo y buscar contenidos concretos [9] o, por el contrario, pueden analizar las preferencias de los usuarios, así como los programas que éstos han visto, con el fin de recomendar contenidos personalizados sin una solicitud explícita previa. Esta última posibilidad pretende reducir la implicación de los usuarios en procesos de búsqueda tediosos. En este artículo, presentamos un sistema recomendador, denominado AVATAR, que explora ambas posibilidades.

A medida que la implantación de la TV digital va avanzando, los recomendadores personalizados van ganando cada vez más interés [8], dando como resultado la aparición de diversas aproximaciones, como pueden ser: sistemas expertos, técnicas bayesianas, métodos basados en contenidos, filtrado colaborativo, árboles de decisión o redes neuronales. Con el fin de conseguir las recomendaciones más adecuadas para el usuario, el sistema recomendador que proponemos combina algunas de las técnicas anteriores con un razonamiento semántico sobre los contenidos televisivos, los perfiles de los usuarios y los registros históricos de visionado de contenidos. Estos mecanismos semánticos necesitan, sin duda, un alto grado de normalización y formalización. En este sentido, debemos indicar que utilizamos el estándar TV-Anytime, que normaliza las descripciones de contenidos genéricos de TV, de instancias concretas de programas y de perfiles de usuario. Constituye, por tanto, un marco adecuado para referenciar, localizar y procesar contenidos.

A semejanza de lo ocurrido con Internet, es de esperar la aparición de un nuevo escenario en el que estén disponibles diversos sistemas recomendadores, compitiendo por ser el preferido de los usuarios. Por tanto, no parece adecuado que el software recomendador se instale durante la fabricación del receptor de TV (Set-Top Box o STB) Por el contrario, proponemos la utilización de aplicaciones descargadas desde el proveedor de servicio a través del flujo de transporte. En el campo de la TV digital en Europa, la adopción del estándar MHP [4] previsiblemente solucionará los problemas actuales de normalización de los equipos receptores, dando lugar a un mercado horizontal para los proveedores de contenidos, proveedores de servicios, plataformas digitales, operadores de red y usuarios. En este modelo de mercado, parece más adecuado el desarrollo de una aplicación independiente de la plataforma que permita realizar recomendaciones no ligadas a intereses comerciales concretos. Es previsible considerar que este modelo requerirá que el coste de las recomendaciones realizadas sea asumido por el usuario, bien directamente o bien a través de publicidad.

Este artículo se organiza de la siguiente forma: la siguiente sección describe la arquitectura del sistema que proponemos. En la sección 3 presentamos el marco conceptual definido para dar soporte al recomendador semántico propuesto, el cual se presenta en la sección 4. La sección 5 presenta un ejemplo del tipo de razonamientos soportados por el recomendador. Finalmente, presentamos algunas conclusiones y líneas de trabajo futuro.

2. Arquitectura del sistema AVATAR

En esta sección presentamos las principales decisiones de diseño de la arquitectura del recomendador AVATAR. El objetivo es que el sistema presente un alto grado de modularidad con el fin de que se puedan añadir fácilmente nuevas funcionalidades relacionadas con la elaboración de recomendaciones personalizadas.

A la hora de diseñar el sistema, nos ha preocupado especialmente el desarrollo eficiente de distintas estrategias

*Este trabajo ha sido subvencionado por el Ministerio de Educación y Ciencia dentro del proyecto TSI2004-03677

para elaborar sugerencias de contenidos televisivos ya que las técnicas actuales tienen una capacidades de razonamiento bastante limitadas de forma individual. Por tanto, proponemos una arquitectura abierta que: (i) permita actualizar los módulos que generan recomendaciones y añadir nuevos módulos que calculen sugerencias siguiendo diferentes estrategias, y (ii) soporte mecanismos para combinar varias recomendaciones. En concreto, en la versión actual del sistema, proponemos una aproximación que combina un modelo bayesiano y técnicas de razonamiento semántico, una metodología de inferencia usual en el campo de la Web Semántica [5, 3], pero novedosa en el campo de los recomendadores televisivos.

La información sobre los usuarios que necesita el recomendador incluye, al menos, datos personales objetivos, sus preferencias en relación con los contenidos televisivos y registros históricos que almacenen los programas vistos en el pasado. Además de los datos proporcionados por los usuarios, AVATAR debe tener en cuenta el éxito obtenido en recomendaciones previas con el fin de mejorar la calidad de las sugerencias futuras. Esta información de realimentación permite conseguir una rápida convergencia entre las preferencias del usuario y las recomendaciones realizadas.

2.1. Una aplicación interactiva MHP

Como comentamos en la introducción, el sistema debe ser lo suficientemente flexible para actualizarse con frecuencia. Por este motivo, proponemos utilizar aplicaciones interactivas descargadas desde el proveedor de servicio hasta el STB a través del flujo de transporte. Actualmente existen importantes esfuerzos para normalizar este tipo de aplicaciones. Distintas organizaciones como DBV (*www.dvb.org*) y ATSC (*www.atsc.org*) han identificado la necesidad de un proceso de normalización para asegurar la interoperabilidad entre aplicaciones de diferentes proveedores y sobre STB de diferentes fabricantes. Esta normalización se ocupa del ciclo de vida de las aplicaciones, los mecanismos de control y del API disponible localmente para las aplicaciones, de forma independiente de los recursos de un receptor concreto.

La iniciativa de normalización que más ha avanzado es DVB MHP, estándar que sigue nuestra propuesta. En este estándar, las aplicaciones se ejecutan ligadas a servicios concretos o eventos (programas de TV) en un servicio, y, normalmente, no sobreviven a la finalización de ese contexto (*cambio de servicio* o *final de evento*) No obstante, en este tipo de recomendadores es necesario un proceso de realimentación persistente para medir el éxito de las sugerencias, almacenando referencias a los programas vistos por los usuarios para analizar el resultado de las recomendaciones realizadas. Esta tarea se extiende más allá de un servicio o programa concreto, por lo que no puede ser realizada por una aplicación MHP. Por tanto, surge la necesidad de un agente especial que almacene de manera ininterrumpida las acciones del usuario. En nuestra propuesta este agente, denominado **agente local** se integra en el software del sistema.

Otro aspecto que necesita normalización es el referente a los formatos de la información que describe la programación de los canales de televisión. Actualmente las diferentes plataformas digitales envían una pequeña cantidad de datos sobre los eventos que van a programar, con el fin de permitir la elaboración de Guías Electrónicas de Programación que muestren los contenidos de TV de una manera organizada. Esta información es insuficiente para procesos complejos, como pueden ser aquellos realizados por recomendadores basados en razonamiento semántico, y, por tanto, se debe ampliar.

Si esta información es enviada por el proveedor de servicios, podría darse el caso de que tenga un formato de almacenamiento propietario, algo no deseable en un escenario en el que estén compitiendo varios servicios de recomendación, ya que podría suponer la existencia de gran cantidad de información redundante sobre la programación. Lo deseable será que esta información se envíe una sola vez en un formato normalizado y esté disponible para todos los proveedores de servicio. Hoy en día existe una iniciativa muy importante promovida por el Foro TV-Anytime (*www.tv-anytime.org*), que pretende normalizar los formatos de los metadatos que describan los contenidos de TV y los perfiles de los usuarios. Las principales organizaciones de normalización (DVB y ATSC) están trabajando en integrar TV-Anytime en sus respectivas infraestructuras.

En nuestro prototipo hemos utilizado los formatos definidos por TV-Anytime para almacenar los registros históricos de los usuarios y sus preferencias personales sobre contenidos televisivos. Estos datos se almacenan en los perfiles de usuario, junto con información adicional, como comentaremos en el apartado 3.2. Toda esta información deberá ser almacenada de manera permanente en el propio STB por parte del agente local. Este hecho revela una necesidad adicional de estandarización: el acceso a esta información debería realizarse de forma normalizada ya que el formato concreto de almacenamiento será desconocido. Este aspecto aún no ha sido abordado, ni por MHP ni por TV-Anytime. Por esta razón, hemos introducido una nueva API MHP – la API MHP TV-Anytime – cuyo propósito es proporcionar una forma independiente y neutral de acceder a la información descrita mediante los metadatos TV-Anytime, aún cuando el agente local no utilice el formato TV-Anytime para almacenar estos datos.

2.2. AVATAR: un sistema multi-agente

Considerando todos los aspectos indicados, proponemos un arquitectura modular y abierta (Fig. 1), en la que el recomendador está dividido en dos partes. La primera está compuesta por el software local del STB que, por una parte solicita los datos personales y las preferencias de los usuarios, y por otra almacena la información relativa a los contenidos de TV ya vistos. Esta información estará accesible al recomendador a través del API MHP TV-Anytime (en formato TV-Anytime) La segunda parte consiste en la aplicación MHP que implementa el servicio de recomendación, y en la que identificamos tres módulos:

Recomendadores. Son varios agentes que implementan las diferentes estrategias para realizar las recomendaciones personalizadas. La figura 1 muestra los tres implementados en la versión actual de nuestro recomendador: un agente basado en técnicas bayesianas, uno basado en razonamiento semántico y otro basado en coincidencia de perfiles[1]. Cada uno de estos agentes realiza recomendaciones de forma independiente (B-REC, S-REC y P-REC) y el módulo de combinación toma como entrada estas tres sugerencias para calcular una recomendación global (G-REC) El agente de combinación es una red neuronal, cuya utilización

[1]En este artículo nos centraremos en la descripción del agente basado en razonamiento semántico. El lector interesado en el diseño de los otros agentes puede consultar [1].

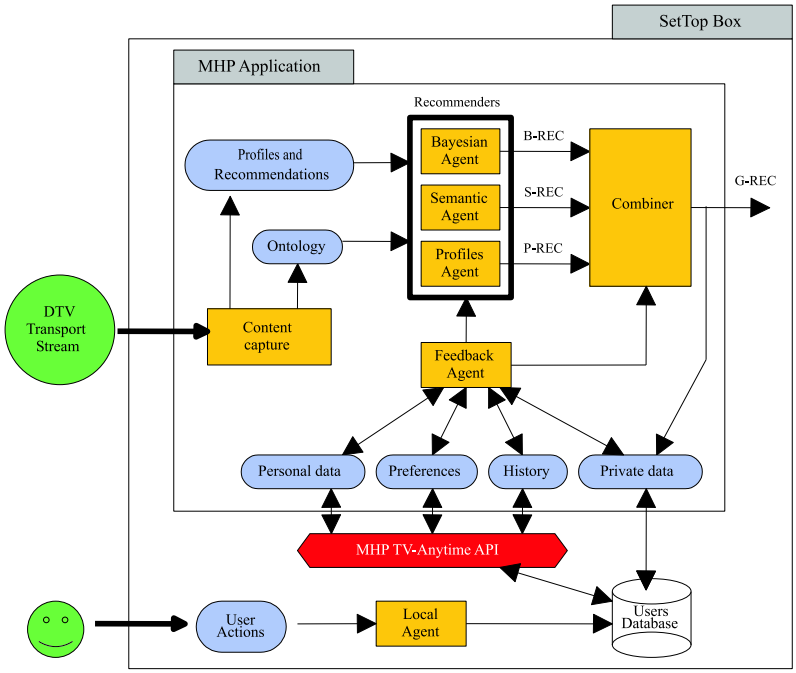


Figura 1: Arquitectura del sistema recomendador

como estrategia de combinación está justificada en diversos trabajos, como puede ser [11]. La recomendación final se almacena en el STB y se compara con las decisiones del usuario para mejorar futuras sugerencias.

Recogida y clasificación de la información. Los algoritmos utilizados por AVATAR necesitan información sobre la programación de los canales de TV y un grupo de perfiles de usuario prototípicos con recomendaciones generales para cada uno de estos perfiles. Este subsistema se encarga de extraer la información comentada de los flujos de transporte y la clasifica adecuadamente para su almacenamiento. En nuestra propuesta, toda la información sobre contenidos de TV se recibe en formato TV-Anytime.

Realimentación. Como ya hemos comentado, la información con la que trabaja el sistema recomendador debería actualizarse a menudo para que las preferencias asociadas a los perfiles de los usuarios sean adecuadas. Para llevar a cabo este proceso, AVATAR necesita acceder a la información proporcionada por los espectadores una vez se les hayan presentado las sugerencias. Para ello el agente de realimentación accede a la información almacenada en la base de datos que gestiona el agente local a través de la API MHP TV-Anytime, y actualiza los perfiles de usuario. Además, el agente de realimentación realimenta el módulo formado por los agentes recomendadores con la información que éstos necesiten, y que dependerá de los métodos que implementen. Así, los agentes semánticos necesitan información actualizada sobre las preferencias de los usuarios y el conocimiento almacenado en la ontología de TV implementada. Los agentes basados en coincidencia de perfiles utilizan perfiles prototípicos recibidos desde el proveedor e información sobre los gustos de los usuarios sobre contenidos de TV. En lo referente a los agentes bayesianos, es necesaria información referente a los programas ya vistos y sobre las preferencias de los usuarios registrados en el sistema.

3. El marco de razonamiento

En esta sección describimos los elementos que integran el marco de razonamiento propuesto en nuestro enfoque. Empezaremos presentando la base de conocimiento del sistema, es decir, la ontología que hemos implementado. A continuación, propondremos un forma de representación de las preferencias de los usuarios basada en la reutilización de dicha ontología. Finalmente, presentamos una clasificación de las diferentes relaciones semánticas que nuestro recomendador puede inferir.

3.1. La ontología televisiva

Para obtener un recomendador de contenidos personalizados basado en razonamiento semántico es necesario representar de forma organizada el conocimiento del dominio de aplicación concreto del sistema, en este caso, la TV. En el marco de la Web Semántica una de las tecnologías de uso más extendido para este tipo de propósitos son las ontologías [10]. Por esta razón, hemos optado por implementar, mediante el lenguaje OWL [6] y utilizando la herramienta Protégé-2000 (*http://protege.stanford.edu*), una ontología sobre el dominio de la TV[2].

Así, hemos definido las diferentes entidades relacionadas con los contenidos audiovisuales que queremos considerar en nuestro análisis, recurriendo para ello a un conjunto de clases. Por ejemplo, en la ontología definimos una clase raíz, *TV Contents*, de la que dependerán las diferentes categorías de programas considerados en nuestro enfoque (*Movies*, *TV Series*, *Informative Programs*, *Sport Programs*, etc.). Del mismo modo, también definiremos clases para identificar a las personas que participan en los programas (*Presenters*, *Directors*, *Actors*, *Actresses*, *Producers*, *ScreenWriters*, etc.)

Las especializaciones relativas a las clases descritas antes, son traducidas a la ontología en forma de subclases (e.g. *Action Movies*, *Terror Series*, *National News*, *Soccer*

[2]Esta ontología está disponible en *http://avatar.det.uvigo.es/ontology*.

Broadcasting, *Movie Directors*, *StarringActors*, etc.)

Las características de los diferentes programas de TV que define la especificación de metadatos de TV-Anytime, han sido introducidos en nuestra base de conocimiento mediante propiedades, también organizadas jerárquicamente. Hemos utilizado las diferentes propiedades que define OWL – funcionales, inversas, simétricas, transitivas, entre otras –, con el objeto de enriquecer la base de conocimiento, y facilitar los procesos de inferencia.

Una vez definida la jerarquía de clases y propiedades de la ontología OWL, se hace necesario instanciar esta base de conocimiento, utilizando para ello programas de TV concretos. En este proceso, se ha analizado la variada oferta de las principales plataformas de TV disponibles, con el fin de obtener una ontología completa a partir de la cual, validar el enfoque de razonamiento propuesto.

3.2. Los perfiles de usuario

Los perfiles de usuario, son uno de los elementos clave en el proceso de personalización, ya que es necesario un mecanismo para representar formalmente el conocimiento relativo a las preferencias de los usuarios. Cada perfil definido en AVATAR almacena información personal sobre el usuario, así como sus preferencias. Para modelar estas preferencias hemos usado un subconjunto dinámico de la ontología OWL de TV descrita antes, construido de forma incremental a medida que el sistema conoce información adicional sobre el espectador. Dicho subconjunto, integrado por propiedades, clases e instancias concretas, recibe el nombre de *contexto semántico personalizado* ó $\mathcal{CSP}$. Por razones obvias, los perfiles de usuario en AVATAR reciben el nombre de *perfiles-ontología*. Concretamente, cuando el usuario ve un nuevo contenido, AVATAR añade a su perfil (i) esa instancia, (ii) la clase a la que ésta pertenece, (iii) la jerarquía completa de superclases relativas a ella en la ontología de TV y (iv) algunas propiedades que permiten identificar características de los programas relevantes en nuestro entorno de personalización (género, créditos del programa, palabras claves, etc.).

Merece la pena destacar que los *perfiles-ontología* superan claramente el empleo de listas planas contemplado en enfoques anteriores. En éstas, se identifican las preferencias de los usuarios pero no se estructuran de forma que se favorezca el descubrimiento de nuevo conocimiento, objetivo claramente perseguido en nuestro recomendador semántico.

Por otra parte, para mantener preferencias de usuario permanentemente actualizadas, nuestra propuesta define tres índices que estarán asociados a cada instancia y a cada clase contenida en el *perfil-ontología* del usuario. Estos índices son actualizados por el agente de realimentación de forma automática a partir de las acciones llevadas a cabo por los espectadores (por ejemplo, qué programas son aceptados y cuáles son rechazados de entre todos los sugeridos por el recomendador), tal como se describe en [2]. Los índices considerados son:

- *Nivel de Interés* ó DOI: Cuantifica el interés o desinterés del usuario en la instancia/clase a la que dicho índice esté asociado. Los valores positivos están reservados para aquellas instancias del *perfil-ontología* que interesan al usuario y a las que hemos denominado preferencias positivas. Por el contrario, las preferencias negativas hacen referencia a aquellas instancias que no agradan al espectador.

- *Confianza*: Este índice se usa para cuantificar el éxito o fracaso del recomendador en las sugerencias ofrecidas en el pasado.

- *Relevancia*: Resulta de la combinación de los dos primeros y es usado por el recomendador para ordenar los contenidos finalmente sugeridos al usuario.

El proceso de razonamiento que proponemos partirá de la información expresada en el $\mathcal{CSP}$ del usuario y, a partir de ella, calculará las recomendaciones oportunas.

3.3. Clasificación de relaciones semánticas

Una vez que disponemos de la información de partida del proceso de razonamiento en los perfiles de usuario, y antes de describir este proceso, es necesario presentar las diferentes relaciones semánticas que **AVATAR** puede inferir a partir de dicha información.

En realidad, el conjunto de información de partida del proceso de razonamiento no será sólo el formado por las clases y propiedades definidas en el $\mathcal{CSP}$ del espectador, sino que estará formado por esas clases y propiedades junto con aquellas relacionadas con ellas desde un punto de vista semántico. Definimos esta relación mediante los siguientes conceptos:

- **Nexo semántico entre propiedades:** Dadas dos propiedades p_i y p_j, decimos que existe un nexo semántico entre ellas (lo que expresaremos como $p_i \leftrightarrow p_j$) si se cumple alguna de las siguientes condiciones:

- $p_i = p_j$

- p_i es la superpropiedad directa de p_j en la jerarquía definida en la ontología[3]. Expresaremos esta condición como $p_j \sqsubseteq_D p_i$.

- p_j es la superpropiedad directa de p_i en la jerarquía.

- $\exists$ q, tal que $p_i, p_j \sqsubseteq_D q$: Es decir, p_i y p_j son propiedades hermanas.

- $\exists\ r$ y q, tales que $p_i \sqsubseteq_D r \sqsubseteq_D q$ y $p_j \sqsubseteq_D q$. Es decir, p_j es propiedad hermana de la superpropiedad directa de p_i.

- $\exists\ r$ y q, tales que $p_j \sqsubseteq_D r \sqsubseteq_D q$ y $p_i \sqsubseteq_D q$. Es decir, p_j es subpropiedad directa de una propiedad hermana de p_i.

- p_i es propiedad inversa de p_j (representaremos la inversa de p_j como $\neg p_j$)

En la notación introducida antes, esta definición sería:

$$
\begin{array}{lll}
p_i \leftrightarrow p_j & \Leftrightarrow & p_i = p_j \ \vee \\
 & & p_j \sqsubseteq_D p_i \ \vee \\
 & & p_i \sqsubseteq_D p_j \ \vee \\
 & & \exists\ \mathrm{q} \,/\, p_i, p_j \sqsubseteq_D q \ \vee \\
 & & \exists\ \mathrm{r,q} \,/\, p_i \sqsubseteq_D r \sqsubseteq_D q \ \wedge \ p_j \sqsubseteq_D \mathrm{q} \ \vee \\
 & & \exists\ \mathrm{r,q} \,/\, p_j \sqsubseteq_D r \sqsubseteq q \ \wedge \ p_i \sqsubseteq_D \mathrm{q} \ \vee \\
 & & p_i = \neg p_j
\end{array}
$$

[3]De todas las propiedades relacionadas mediante *subsumption* con p_i (representado como *Properties* $\sqsubseteq p_i$), únicamente consideraremos la(s) superpropiedad(es) directa(s). Para identificar esta relación utilizaremos el operador $\sqsubseteq_D$.

- **Nexo semántico entre clases:** Análogamente, definimos un nexo semántico entre las clases C_i y C_j como:

$$
\begin{aligned}
C_i \leftrightarrow C_j \Leftrightarrow\ & C_i = C_j \ \vee \\
& C_j \sqsubseteq_D C_i \ \vee \\
& C_i \sqsubseteq_D C_j \ \vee \\
& \exists\, Q^* \neq \textit{Ontology Root} \ / \ C_i, C_j \sqsubseteq_D Q^* \ \vee \\
& \exists\, A, B \,/\, C_i \sqsubseteq A \sqsubseteq_D B \ \wedge \ C_j \sqsubseteq B \ \vee \\
& \exists\, A, B \,/\, C_j \sqsubseteq_D A \sqsubseteq_D B \ \wedge \ C_i \sqsubseteq_D B
\end{aligned}
$$

En la definición anterior, $\mathbf{Q}^*$ debe ser necesariamente diferente del elemento raíz de la ontología. De lo contrario, estaríamos relacionando contenidos muy diferentes entre sí, solamente por ser subclases directas del elemento raíz.

Una vez que hemos establecido los criterios para descubrir relaciones entre clases y propiedades, nos centramos en las asociaciones establecidas entre las diferentes instancias declaradas en la ontología del sistema.

Empecemos considerando que dos instancias de clases definidas en una ontología pueden estar relacionadas **explícitamente** a través de propiedades *Object* declaradas en la base de conocimiento, o bien puede haber **relaciones implícitas** entre ellas, a través de instancias de otras clases. Para referirnos a relaciones semánticas *explícitas*, emplearemos la notación C_1 [p] C_2, siendo p una propiedad *Object*, y C_1 y C_2 las clases a las que pertenecen las instancias entre las que se establece la relación.

El descubrimiento de relaciones *implícitas* será el objetivo del recomendador **AVATAR** durante el proceso de razonamiento semántico, ya que permitirá obtener mejores recomendaciones con un alto grado de personalización. En concreto, en nuestra propuesta definimos los siguientes tipos de relaciones semánticas *implícitas*:

- **Cadena semántica directa** ($\Rightarrow$):Está formado por varias relaciones semánticas *explícitas* de la forma:

  D_0 $[p_0]$ D_1 $[p_1]$ D_2 ... D_{q-1} $[p_{q-1}]$ D_q

  Este tipo de relación se basa en el principio de transitividad (si $a \rightarrow b$ y $b \rightarrow c$ entonces $a \rightarrow c$), y permite una relación semántica entre instancias de las clases D_0 y D_q. Lo denotaremos como $D_0 \Rightarrow D_q$.

- **Cadenas semánticas directas con nodos comunes** ($\Rightarrow_{Nodes}$): Consideremos dos cadenas semánticas directas, tales que en ambos participan instancias pertenecientes a una misma clase (en adelante, a este conjunto de clases comunes les llamaremos *Nodos*). Por ejemplo,

  $\mathcal{DSC}_1 = E_0$ $[p_0]$ E_1 $[p_1]$ **N** ... **M** $[p_{m-1}]$ E_m

  $\mathcal{DSC}_2 = D_0$ $[q_0]$ **M** $[q_1]$ D_2 ... **N** $[q_{n-1}]$ D_n

  $DSC_1 \sqcap DSC_2$ = {*Nodos*} = {**M**,**N**}

  El sistema puede inferir las relaciones semánticas $E_0 \Rightarrow E_m$ y $D_0 \Rightarrow D_n$ debido a la existencia de dos *cadenas semánticas directas*. Además por el hecho de haber nodos comunes entre ambas cadenas (M y N), también se puede establecer una relación semántica entre E_0 y D_0. Esta relación es de tipo *cadenas semánticas directas con nodos comunes*, que denotaremos como $E_0 \Rightarrow_{M-N} D_0$.

- **Cadenas semánticas directas mixtas** ($\Rightarrow_{\Rightarrow_{Nodes}}$): Es una combinación de las dos relaciones semánticas anteriores. Así, si una instancia de la clase C_1 está relacionada con otra de la clase C_2 mediante una cadena semántica directa con nodos comunes ($C_1 \Rightarrow_{Nodes} C_2$), y a la vez C_2 está relacionada con otra instancia de C_3 mediante una cadena semántica directa ($C_2 \Rightarrow C_3$), podemos establecer una nueva relación entre C_1 y C_3 gracias al principio de transitividad. A este tipo de asociación semántica la denominaremos *cadena semántica directa mixta*, que denotaremos como $C_1 \Rightarrow_{\Rightarrow_{Nodes}} C_3$.

- **Cadenas semánticas indirectas** ($\approx$): Dadas dos *cadenas semánticas directas* DSC_1 y DSC_2, de longitudes l_1 y l_2, respectivamente, con $l_1 \leq l_2$:

  $\mathcal{DSC}_1 = E_0$ $[p_0]$ E_1 $[p_1]$ E_2 ... $[p_{l_1-2}]$ E_{l_1-1}

  $\mathcal{DSC}_2 = F_0$ $[q_0]$ F_1 $[q_1]$ F_2 ... $[q_{l_2-2}]$ F_{l_2-1}

  Decimos que existe una *cadena semántica indirecta* entre las instancias de las clases E_0 y F_0, si las propiedades involucradas en $\mathcal{DSC}_1$ y $\mathcal{DSC}_2$, están relacionadas una a una mediante *nexo semántico*. Si las longitudes de ambas cadenas puede son diferentes, la condición debe verificarse hasta que se termine la cadena más corta de ambas. Es decir, para las dos *cadenas semánticas directas* presentadas arriba, se cumplirá $E_0 \approx F_0 \Leftrightarrow p_i \leftrightarrow q_i \ \forall$ i (0 $\leq$ i $\leq l_1 - 2$)

- **Cadenas semánticas virtuales** ($\approx_{\Rightarrow}$, $\approx_{\Rightarrow_{Nodes}}$): Las *cadenas semánticas indirectas* permiten inferir otro tipo de asociaciones en la base de conocimiento del sistema. Estas nuevas asociaciones se pueden clasificar en dos tipos. El primero, que denominaremos *cadena semántica virtual de tipo I*, se obtendrá combinando *cadenas semánticas directas* y *cadenas semánticas indirectas* mediante el principio de transitividad (así por ejemplo, si $C_1 \approx D_1$ y $D_1 \Rightarrow D_3$, entonces $C_1 \approx_{\Rightarrow} D_3$) De forma similar, podemos obtener *cadenas semánticas virtuales de tipo II* combinando *cadenas semánticas directas con nodos comunes* y *cadenas semánticas indirectas*. Los operadores de ambos tipos de *cadenas virtuales* son $\approx_{\Rightarrow}$ y $\approx_{\Rightarrow_{Nodes}}$ respectivamente.

  En la figura 2 se muestran algunos ejemplos de las relaciones semánticas definidas.

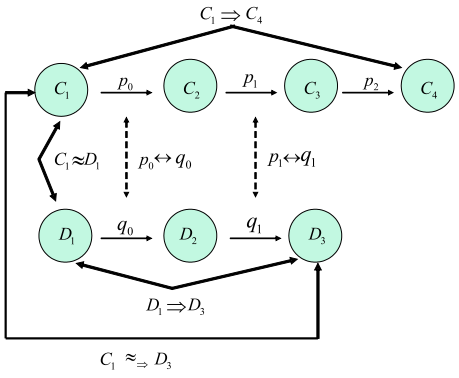


Figura 2: Relaciones semánticas inferidas

4. Inferencia de relaciones semánticas

En esta sección presentamos el proceso que llevará a cabo el agente recomendador semántico, dentro del mar-

co presentado antes, para proporcionar una recomendación adecuada y personalizada a un usuario.

El algoritmo que ejecutará este agente (Algoritmo 1) recibe como entrada el $\mathcal{CSP}$ del usuario y devuelve una relación priorizada de programas como recomendación final del sistema al usuario.

En primer lugar, será necesario calcular las clases y propiedades de interés en el proceso de razonamiento semántico. Para ello el algoritmo recorre las clases y propiedades definidas en el $\mathcal{CSP}$ del usuario, y calcula los conjuntos de *clases objetivo* (TC_{C_i}) (que contendrá las clases que estén relacionadas mediante *nexo semántico* con la clase C_i definida en el $\mathcal{CSP}$ del usuario) y de *propiedades objetivo* (TP_{C_i}) (que contendrá las propiedades relacionadas con el conjunto de propiedades asociado a la clase C_i en el $\mathcal{CSP}$ del usuario) A continuación, es necesario recorrer tanto las instancias definidas en cada una de las clases del $\mathcal{CSP}$ del usuario, como las instancias de las clases definidas en el conjunto TC_{C_i} calculado (en el código identificadas como *SetOfInstances*). Con cada una de dichas instancias ($Prog_{C_j}$ [k] en el código) se invoca la función *Direct_Semantic_Chains*, para calcular todas las cadenas semánticas directas relevantes. A partir de este conjunto de cadenas semánticas se pueden descubrir el resto de asociaciones semánticas. Para ello se invocan las funciones correspondientes con los parámetros adecuados.

Una vez descubiertas las asociaciones semánticas, será necesario elaborar la recomendación personalizada. Para ello, el sistema utiliza la función *Recommend*. Esta función clasifica las relaciones descubiertas en base a una estimación de su relevancia. Así por ejemplo, asigna un peso más alto a aquellas relaciones en que participan valores concretos y clases incluidas en el perfil del usuario.

Algoritmo 1

do_Semantic_Recommendation ($\mathcal{CSP}$){
  for each class $C_j \in \mathcal{CSP}$ {
    TC_{C_j} = *Target_Classes* (C_j)
    TP_{C_j} = *Target_Properties* (R_{C_j})
    for (k = 0 to |*SetOfInstances*|)
      $\mathcal{DSC}$ = *Direct_Semantic_Chains* ($Prog_{C_j}[k]$)
  }
  for (j = 0 to | $\mathcal{DSC}$ |){
    for (k = j+1 to | $\mathcal{DSC}$ |){
      Rel = *Direct_Semantic_Chains_Common_Nodes*($\mathcal{DSC}_j$,$\mathcal{DSC}_k$)
      $\mathcal{DSCCN}$ = $\mathcal{DSCCN} \sqcup$ *Rel*
      $\mathcal{ISC}$ = $\mathcal{ISC} \sqcup$ *Indirect_Semantic_Chains*($\mathcal{DSC}_j$, $\mathcal{DSC}_k$)
    }
  }
  $\mathcal{MDSC}$ = *Mixed_Direct_Semantic_Chains* ($\mathcal{DSC}$, $\mathcal{DSCCN}$)
  $\mathcal{VSC}$ = *Virtual_Semantic_Chains* ($\mathcal{DSC}$, $\mathcal{DSCCN}$, $\mathcal{ISC}$)
  Recommend ($\mathcal{DSC}$, $\mathcal{DSCCN}$, $\mathcal{MDSC}$, $\mathcal{ISC}$, $\mathcal{VSC}$)
}

5. Un ejemplo de recomendación

Asumamos un *perfil-ontología* que contiene los siguientes programas: (i) una retransmisión de Formula 1 donde interviene el piloto Fernando Alonso, y (ii) una película de acción, cuya trama discurre en Mexico, protagonizada por Clint Eastwood. A partir de esta descripción es fácil ver que las clases del $\mathcal{CSP}$ de este perfil constará de las siguientes clases y propiedades, representadas mediante C_i y R_{C_i} (i=0,1), respectivamente:

C_0 = Formula 1 Broadcasting
R_{C_0} = {hasDrivers}
C_1 = Action Movies
R_{C_1} = {hasPlace, hasStarringActor, hasName}

Supongamos que nuestra base de datos contiene, entre otros, los siguientes programas de TV como instancias de clases de la ontología de TV:

- Una comedia (clase *Comedy Movies*) en la que "Clint Eastwood", un actor que puede interesar al usuario, participa como actor secundario.

- Un concurso de TV (clase *Quiz Shows*) presentado por una actriz de telenovelas (clase *Soap Operas*), que ha protagonizado la comedia anterior.

- Una película basada en los asesinatos de hinchas de un equipo de fútbol (instancia de la clase *Terror Movies*) cuya trama transcurre en el estadio de fútbol "Amsterdam Arena", producida por "Clint Eastwood".

- Un partido de fútbol del mundial 2002 (clase *Soccer Broadcastings*), en el que jugaban, entre otros, el futbolista "Michael Owen", celebrado en el estadio "Yokohama".

- Un documental deportivo (clase *Sport Documentaries*) sobre los deportistas de élite, con participaciones del piloto "Fernando Alonso" y del futbolista "Beckham".

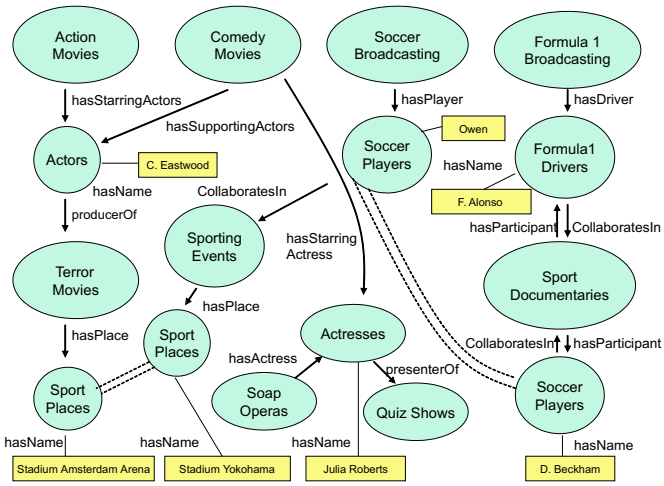


Figura 3: Subconjunto de las instancias de la ontología

En la figura 3 se muestra una parte de la ontología de TV implementada[4], que contiene los programas que acabamos de definir. En realidad, las clases del ejemplo tienen muchas más propiedades de las mostradas, pero el subconjunto mostrado es suficiente para ilustrar el ejemplo que nos ocupa. Esto también nos permite demostrar que el sistema de recomendación es capaz de inferir relaciones semánticas en la base de conocimiento, aun cuando la información proporcionada sea incompleta o escasa. Como se aprecia en la figura, emplearemos rectángulos para definir los valores tomados por las propiedades *Datatype*.

Veamos en qué medida son importante los programas que hemos representado en la figura 3, de cara a una recomendación personalizada. Tal como se ha descrito en el

[4]En esta figura, mostramos únicamente un subconjunto reducido de instancias. La jerarquía de clases se ha omitido por falta de espacio.

algoritmo *do_Semantic_Recommendation*, el primer paso será calcular el conjunto de clases y propiedades de interés a la hora de razonar (*clases objetivo* y *propiedades objetivo*), a partir de la información del $\mathcal{CSP}$ del usuario. Así, la función *Target_Classes*, permite obtener las clases relevantes para el proceso de recomendación, asociadas a las clases C_0 y C_1 del $\mathcal{CSP}$. En nuestro ejemplo, éstas serían:

- $\mathcal{TC}_{C_0}$, formado por *Formula 1 Broadcasting* y *Soccer Broadcasting* (clase hermana de C_0)
- $\mathcal{TC}_{C_1}$, formado por *Action Movies* y *Soap Operas* (hermana de la superclase de C_1)

En el ejemplo mostrado en la figura 3, todas las propiedades definidas son relevantes en el proceso de razonamiento semántico (por pertenecer a $\mathcal{TP}_{C_0}$ ó $\mathcal{TP}_{C_1}$)

Una vez calculados los conjuntos de clases y propiedades en las que se basará el sistema para razonar, podemos empezar a describir las relaciones inferidas por el sistema, y las razones por las que el sistema incluiría los contenidos de la figura 3 en la recomendación final.

En primer lugar, tenemos una relación entre la retransmisión de Fórmula 1 vista por el usuario, y un documental deportivo en el que participan tanto el piloto "Fernando Alonso" (presente en el perfil del usuario), como el futbolista "Beckham". En la figura 3 también se aprecia una relación entre una retransmisión de fútbol (instancia de la clase *Soccer Broadcasting*) y el estadio "Yokohama", donde éste tiene lugar. Además en esta última cadena semántica directa también participa una instancia de la clase *Soccer Players*, por lo que el sistema descubre un nexo común entre el documental deportivo descubierto antes, y esta retransmisión de fútbol. Esta relación de tipo *cadena semántica directa con nodos comunes*, permite que el sistema infiera una asociación entre ambos programas, aún cuando no es el mismo futbolista en el que interviene en ambos. Este tipo de inferencias sólo tiene cabida en un entorno de razonamiento sobre la semántica de los contenidos de TV. En resumen, a partir de $\mathcal{TC}_{C_0}$, el sistema infiere, entre otras, las siguientes relaciones:

Formula 1 Broadcasting $\Rightarrow$ *Sport Documentaries*
Soccer Broadcasting $\Rightarrow$ *Sport Places*
Sport Documentaries $\Rightarrow_{\textit{Soccer Players}}$ *Soccer Broadcasting*
Formula 1 Broadcasting $\Rightarrow\Rightarrow_{\textit{Soccer Players}}$ *Sport Soccer Broadcasting*

De los dos nuevos contenidos de TV descubiertos por el sistema (*soccer broadcasting* y *sport documentaries*), el segundo es especialmente relevante, ya que está relacionado con uno de los valores presentes en el perfil de usuario: el piloto de Fórmula 1. Sin embargo, el sistema también debe considerar en el proceso de recomendación, la importancia del programa sobre fútbol descubierto, debido a que este tipo de programas y las retransmisiones de Fórmula 1 (vistas por el usuario) son instancias de clases hermanas.

Sigamos descubriendo otras relaciones semánticas existentes en la figura 3. Podemos ver que hay una *cadena semántica directa* entre la película de acción vista por el usuario, y una película de terror, establecida a través del actor "Clint Eastwood". Además esta película está relacionada con la instancia "Amsterdam Arena" de la clase *Sport Places*, ya que la trama de la película transcurre en este estadio. Esto permite descubrir un nexo de unión entre la misma y el partido de fútbol inferido anteriormente: ambas instancias están asociadas a lugares donde tienen lugar eventos deportivos (clase "Sport Places"), y por tanto hay una relación del tipo *cadenas semánticas directas con nodos comunes* entre ambas.

La película de terror se incluirá en la recomendación final por los dos motivos que acabamos de presentar. Por una parte, interviene en ella el actor definido en el perfil de usuario, y por otro lado, está relacionado semánticamente con el partido de fútbol, cuya relevancia en el proceso de personalización ya ha sido explicada antes.

Las relaciones que hemos descrito se resumen así:

Action Movies $\Rightarrow$ *Terror Movies*
Terror Movies $\Rightarrow$ *Sport Places*
Terror Movies $\Rightarrow_{\textit{Sport Places}}$ *Soccer Broadcasting*

Exploremos ahora las restantes clases del conjunto $\mathcal{TC}_{C_1}$. Por ejemplo, podemos ver en la figura 3, una *cadena semántica directa* entre la telenovela y el concurso. Además dicha *cadena semántica directa*, permite descubrir sendas *cadenas indirectas*. Veamos cómo: partimos de las siguientes *cadenas semánticas directas*:

Soap Operas [hasActress] *Actresses* [presenterOf] *Quiz Shows*
Action Movies [hasStarringActors] *Actors* [producerOf] *Terror Movies*

Con arreglo a la jerarquía de clases y propiedades establecida en la ontología, la propiedad *hasActress* está relacionada mediante *nexo semántico* con *hasStarringActors*, ya que la primera es hermana de la superpropiedad de la segunda, tal como se muestra en la figura 4.

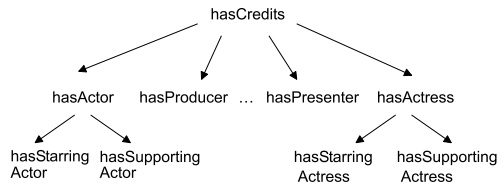


Figura 4: Subconjunto de la jerarquía de propiedades

Dado que existe un *nexo semántico* entre las propiedades *presenterOf* y *producerOf* (son propiedades hermanas), el sistema puede inferir una *relación indirecta* entre las instancias de las clases *Soap Operas* y *Action Movies*. La segunda relación indirecta mencionada antes se establecería de manera análoga entre las instancias de las clases *Soap Operas* y *Comedy Movies*. Es decir, se infieren:

Soap Operas $\approx$ *Action Movies*
Soap Operas $\approx$ *Comedy Movies*

Considerando estas relaciones, la película de acción vista por el usuario estaría relacionada semánticamente con una telenovela. En principio, esta relación indirecta no involucra ningún valor concreto definido en el perfil personal del espectador, por lo que la telenovela no sería especialmente relevante de cara a la recomendación final. Sin embargo, si seguimos razonando sobre la ontología de la figura 3 vemos que es posible descubrir una asociación más directa con valores conocidos por el sistema.

Para ello, centrémonos en la instancia de la clase *Comedy Movies*. Este programa se puede relacionar semánticamente con la película de acción mencionada antes, ya que también en ella participa el actor favorito del usuario.

Además, podemos inferir una nueva cadena semántica entre la instancia de *Comedy Movies* y un concurso televisivo presentado por la actriz protagonista de la comedia. En la figura 3, es fácil ver que también se puede establecer una asociación semántica entre la telenovela y esta comedia, ya que ambas tienen un nexo en común: "Julia Roberts". Estas últimas relaciones descubiertas se resumen como:

Comedy Movies $\Rightarrow_{Actors}$ *Action Movies*
Comedy Movies $\Rightarrow$ *Quiz Shows*
Soap Operas $\Rightarrow$ *Quiz Shows*
Comedy Movies $\Rightarrow_{Actresses}$ *Soap Operas*

A la vista de los contenidos sugeridos por **AVATAR** en este ejemplo, podemos concluir que este tipo de mecanismos de recomendación, basados en el establecimiento de asociaciones semánticas, son especialmente útiles para descubrir al usuario contenidos que no conocía. Estos programas, sin embargo, pueden ser de su interés, debido a que guardan relación, desde un punto de vista semántico, con las preferencias declaradas en su perfil personal.

6. Conclusiones y trabajo futuro

En este artículo se ha presentado el marco de razonamiento semántico de **AVATAR**, un sistema recomendador de contenidos de televisión en el que estamos trabajando, y cuyo principal objetivo es conseguir mejorar las recomendaciones personalizadas presentadas a los espectadores, utilizando para ello la experiencia de la Web Semántica. En dicho marco hemos incluido la implementación de una ontología OWL para representar el conocimiento del dominio de la TV, necesaria para soportar los procesos de inferencia. Además, hemos propuesto un mecanismo de representación de las preferencias de los usuarios, basado en la reutilización de la jerarquía de clases y propiedades, así como instancias específicas, definidas en dicha ontología. Así, nuestro *perfiles-ontologiá* pueden ser construidos incrementalmente a medida que el sistema recibe información adicional sobre las preferencias de los usuarios.

Sobre este marco se propone un algoritmo de razonamiento semántico que permite inferir relaciones entre programas a partir de la información almacenada en la ontología, con el objetivo de conseguir mejores recomendaciones que las proporcionadas por los recomendadores actuales basados en técnicas sintácticas.

Para validar el algoritmo de inferencia que proponemos hemos desarrollado un prototipo del sistema. Para completar este prototipo, debemos abordar como trabajo futuro el problema de ordenar las relaciones descubiertas en base a su relevancia semántica y a la información almacenada en el perfil de usuario. Esto se debe a que la capacidad de razonamiento del sistema **AVATAR**, permite inferir una gran cantidad de relaciones, pero no todas ellas son igual de importantes de cara a la recomendación. En este momento, nuestro trabajo se centra en este proceso de "ponderación". Para ello tendremos en cuenta factores como: (i) el tipo de relación semántica: así por ejemplo, las *cadenas semánticas directas* son relaciones más relevantes que las *indirectas*; (ii) la longitud de las cadenas semánticas, ya que las cadenas que involucran a muchas clases son relaciones menos directas y, por tanto, menos interesantes; (iii) la presencia de valores definidos en el perfil de usuario en las cadenas descubiertas; y (iv) la presencia de campos definidos en la especificación TV-Anytime que pueden ser relevantes a la hora de ofrecer una recomendación: por ejemplo, el campo *Intended Audience* que identifica el público al que va dirigido el programa.

Referencias

[1] Blanco Fernández Y., Pazos Arias J. J., Gil Solla A., Ramos Cabrer M., Barragáns Martínez B. and López Nores M. A Multi-Agent Open Architecture for a TV Recommender System: A Case Study using a Bayesian Strategy. In *Proc. of the IEEE Sixth International Symposium on Multimedia Software Engineering*, 2004.

[2] Blanco Fernández Y., Pazos Arias J. J., Gil Solla A., Ramos Cabrer M., López Nores M. and Barragáns Martínez B. AVATAR: Modeling Users by Dynamic Ontologies in a TV Recommender System based on Semantic Reasoning. In *Proc. of the 3rd European Conference on Interactive Television: User Centred ITV Systems, Programmes and Applications (EuroITV-05)*, 2005.

[3] Daconta M., Obrst L. J. and Smith K. J. *The Semantic Web: A Guide to the Future of XML, Web Services and Knowledge Management*. John Wiley & Sons, 2003.

[4] ETSI TS 101 812. *Multimedia Home Platform (MHP)*. 2002.

[5] Geroimenko V. and Chen C. *Visualizing the Semantic Web*. Springer Verlag, 2003.

[6] McGuinness D. and van Harmelen F. OWL Web Ontology Language Overview. W3C Recommendation. 2004.

[7] Pagani M. *Multimedia and Interactive Digital TV: Managing the Opportunities Created by Digital Convergence*. Idea Group Publishing, 2003.

[8] Ricci F., Arslan B., Mirzadeh N. and Venturini A. ITR: a Case-Based Travel Advisory System. In *Proc. of 6th European Conference on Case Based Reasoning (ECCBR-02)*, 2002.

[9] Schonfeld E. Don't Just Sit There. Do Something. *Business 2.0*, 2000.

[10] Staab S. and Studer R. *Handbook on Ontologies*. Springer, 2004.

[11] Zimmerman J., Kurapati K., Buczak A. L., Schaffer D., Gutta S. and Martino J. TV Personalization System: Design of a TV Show Recommender Engine and Interface. *Personalized Digital Television: Targeting Programs to Individual Viewers*, 2004.

Ontologías para la Medida de la Calidad de Servicio Percibida

Alfonso Sánchez-Macián[1], Luis Bellido[2], Encarna Pastor[2],
[1]Departamento de Ingeniería Informática, Universidad Antonio de Nebrija
C/ Pirineos 55 28040 - Madrid
E-mail: asanchep@nebrija.es
[2]Departamento de Ingeniería de Sistemas Telemáticos, Universidad Politécnica de Madrid
ETSI de Telecomunicación, Ciudad Universitaria s/n. 28040 - Madrid
{lbt, encarna}@dit.upm.es

Abstract. *Quality of service efforts in data services have been traditionally focused on metrics and provision using objective parameters such as availability, bandwidth or delay. However, user perception of quality sometimes differs from the quality measured by service providers. The problem is a lack of understanding of the relationship between objective parameters and perceived quality. This paper focuses on a methodology to measure user perception of quality. The goal is to provide the basic platform and tools that will allow further research on the definition of utility functions mapping objective parameters to perceived quality. The proposed solution is based on a new ontology that will allow to represent and collect user perception of quality and correlate this information with service parameter measurements. The integration of this ontology and other quality of service ontologies in a platform to gather user opinion is analysed.*

1 Introducción

Los servicios en los sistemas informáticos y, en especial, en las redes de computadores, han evolucionado en cuanto a su concepción de la calidad, hacia un modelo basado en la búsqueda de métodos de provisión y de medida de dicha calidad.

Los esfuerzos se han centrado en medir parámetros objetivos del servicio y definir el valor de la calidad en función del valor de dichos parámetros. Sin embargo, se ha prestado menor atención a la percepción que los usuarios tienen de la calidad de los servicios, así como las expectativas que tienen sobre ellos.

La percepción de la calidad es un indicador de la satisfacción del usuario, y por tanto, su medida se considera un factor clave en las estrategias de las empresas proveedoras de servicios. Desde el punto de vista de la ingeniería, el objetivo es ser capaz de ajustar los parámetros objetivos del servicio para obtener una percepción de calidad determinada. El artículo se enmarca dentro de los trabajos realizados para obtener un sistema que permita establecer la relación entre los parámetros objetivos y la calidad percibida.

Esta correspondencia entre calidad percibida y calidad de la red, se representaría mediante funciones de utilidad que permitan conocer cómo cambia la percepción de los clientes sobre la calidad de un servicio a partir de una variación producida en alguno de los parámetros objetivos y qué acciones se podrían tomar para corregir pérdidas en el nivel de calidad. Estas proyecciones del plano subjetivo al objetivo dependen de cada servicio.

Para captar información relativa a percepción de calidad de servicio por parte de los usuarios, es necesario acudir a métodos disponibles en otros ámbitos de conocimiento, basados en encuestas. La automatización de estos métodos facilitaría la recogida de los datos y su procesado y permitiría obtener una muestra periódica de los mismos.

El uso de ontologías va a facilitar esta automatización debido a que proporcionan una definición entendible por humanos y máquinas sobre conceptos de un dominio de conocimiento. Estas definiciones se implementan mediante lenguajes específicos. Así, es posible establecer una comunicación y un almacenamiento basado en los conceptos definidos con la seguridad de que todos los implicados coinciden en su significado.

Este artículo propone un sistema para recopilar los datos de opinión de los usuarios acerca de la percepción de la calidad de los servicios que utilizan. Para ello se ha definido una ontología de calidad de servicio que da soporte a los parámetros de calidad percibida y a las herramientas utilizadas para su obtención. Además se describen las arquitecturas que facilitarán el intercambio de dichos datos.

El texto comienza señalando las herramientas que existen en otros campos para conseguir una medida de la calidad de servicio percibida. Posteriormente se describe una ontología que se usará para recopilar la información de calidad de servicio percibida utilizando una de las técnicas referidas. Seguidamente se realiza una comparativa entre esta ontología y otros estudios existentes. Por último se sugieren varias arquitecturas para aplicar la ontología en un sistema que sirva para medir la calidad de servicio percibida por los usuarios.

2 Calidad de Servicio Percibida

Se han desarrollado diferentes modelos de provisión de calidad y de métricas de calidad. Muchos de estos trabajos se orientan a definir una clasificación del nivel de calidad en función de valores de parámetros objetivos del servicio. En estos planteamientos se ha obviado un parámetro importante: la percepción de la calidad del servicio por parte de los usuarios.

Aunque la calidad proporcionada por los proveedores puede llegar a satisfacer los requisitos del servicio, en muchos casos existe un gap entre la calidad entregada y la que el usuario percibe [22]. Por tanto, es importante detectar qué factores están implicados en la calidad de servicio percibida [1][2][3][4][5].

La percepción y expectativas del usuario son términos subjetivos que dependen de aspectos relacionados con el usuario. Para conseguir resultados correctos es necesario buscar técnicas diseñadas en otros campos, como las estudiadas para satisfacción del cliente en el ámbito de la economía.

ServQual[22] es una de estas metodologías. Se basa en cuestionarios que recogen la percepción de los usuarios de un servicio y sus expectativas. El instrumento ServQual se compone de 22 preguntas relacionadas con las cinco dimensiones que se han considerado importantes para todos los servicios:

- Elementos tangibles: Apariencia de las instalaciones, equipo y personal y material de comunicación.
- Fiabilidad: Habilidad para cumplir con el servicio prometido de una forma adecuada y constante.
- Respuesta: Deseo y capacidad de ayudar a los clientes y proporcionar el servicio de forma rápida.
- Garantía: Conocimiento y cortesía de los empleados y su habilidad para dar confianza e inspirar credibilidad.
- Empatía: El cuidado o atención individualizada que la empresa da a sus consumidores.

Las veintidós cuestiones se presentan en una doble escala para comparar qué debería proporcionar el servicio (expectativas) y sus propiedades actuales desde el punto de vista del cliente.

ServQual ha sido ampliamente criticado por dos razones:

- Las cinco dimensiones que define no se pueden aplicar a todos los servicios. Puede ser necesario añadir alguna nueva dimensión específica a un servicio particular. Se han realizado estudios relacionados con su aplicación a los sistemas de información (ver [6], [11], [16], por ejemplo) que sugieren nuevas dimensiones y cuestiones.
- Algún autor indica que los valores estadísticos obtenidos como una diferencia entre otros dos valores (expectativas menos percepción) no son una buena técnica de medida, puesto que no tienen una fiabilidad lo suficientemente buena. A pesar de esta circunstancia, se podría usar aún ServQual si se consideran ambas escalas de forma independiente.

Aparte de ServQual, existen otras herramientas disponibles para la medida de la calidad percibida como ServPerf o ServImperf que también utilizan cuestionarios para la recopilación de los resultados.

3 Requisitos de Ontologías de Calidad de Servicio

Algunos autores han tratado de definir las características que deben satisfacer las ontologías de calidad de servicio, así como los requisitos previos necesarios para que sea posible el desarrollo de esas ontologías.

[19] considera que una ontología de métrica de calidad de servicio debe estar soportada por otras ontologías más específicas. Define las siguientes:

- Ontología de unidades de medida. Con unidades base, múltiplos, unidades derivadas (como combinación de otras) y sinónimos.
- Ontología de unidades de moneda. Nombres de moneda, múltiplos, símbolos y sinónimos.
- Ontología de propiedades medidas. Propiedades del servicio que se han de medir para poder calcular el valor actual de la calidad.
- Ontología de métodos de medida. Formas de medir las diferentes propiedades.

Desde el punto de vista de la calidad de servicio percibida, en la que se va a utilizar para medir un método basado en encuestas, las dos primeras ontologías no son necesarias. Sin embargo, con la tercera ontología se pueden expresar las propiedades del servicio según ServQual (las cinco dimensiones). Dentro de la última ontología se podría diferenciar entre métodos subjetivos y objetivos de medida, englobando a ServQual dentro de los primeros.

Por otro lado, [13] presenta dos opciones posibles para la definición de parámetros de calidad de servicio para una ontología de calidad de servicio:

- Una ontología formada por términos fácilmente entendibles y la semántica de relación entre ellos.

- Ontologías constructivas con conceptos bien definidos y operadores de composición. Los nuevos parámetros se definirían a partir de los básicos usando los operadores. Por ejemplo, se podría crear el término "tiempo de respuesta medio" a partir de "tiempo de respuesta" y el operador "media".

Para la ontología buscada se plantea la primera de las opciones, debido a que el carácter subjetivo de los parámetros de calidad percibida dificulta la existencia de relaciones entre ellos que se puedan expresar en forma de operadores.

4 Desarrollo de la Ontología

Para la creación de la ontología de calidad de servicio percibida se ha decidido dividir los conceptos en tres ontologías distintas:

- *Ontología de cuestionarios*. Es una ontología de apoyo. Servirá de base a la herramienta ServQual. Podrá ser reutilizada para cualquier aplicación basada en encuestas.

- *Ontología de Calidad de Servicio*. Define los conceptos básicos que se requieren en el campo de la calidad de los servicios.

- *Ontología ServQual*. Define conceptos heredados de la ontología de calidad de servicio percibida, así como instancias de elementos de dicha ontología para el uso del método ServQual. Establece la relación con la ontología de cuestionarios.

A continuación se estudiarán más en detalle cada una de ellas.

4.1 Ontología de Cuestionarios

Se ha escrito como soporte al modelo ServQual. Los principales conceptos que agrupa son:

- Cuestionario: Define el cuestionario completo, permite indicar fecha de publicación y está compuesto por varios grupos de preguntas.

- Grupo de preguntas: Representa divisiones de preguntas relacionadas.

- Pregunta: Cada pregunta del cuestionario tiene un enunciado y una serie de características, entre las que destaca el tipo o el orden que ocupa en el formulario.

- Método de la pregunta: Existen diferentes métodos de preguntar por una información dependiendo de cómo se pretendan recopilar los datos del usuario. Puede ser una pregunta con respuesta abierta, para dar puntuación, para distribuir un valor entre distintas opciones o para seleccionar opciones por preferencia, por ejemplo. Se incluyen cinco instancias.

- Tipo de pregunta: Basado en los diferentes métodos, existen en las encuestas varios tipos de preguntas (diferencial semántica, por ejemplo). Se incluyen cinco instancias.

- Respuesta a cuestionario. Representa las respuestas a un cuestionario. Tendrá como atributos la fecha de publicación de las respuestas, el cuestionario al que responde y las respuestas organizadas en grupos, entre otros.

- Grupo de respuestas. Agrupa respuestas que tienen unas características comunes.

- Respuesta. Concepto de respuesta a una pregunta. Está relacionada con un tipo de pregunta e incluye el texto de la respuesta y el orden en el formulario, entre otras propiedades.

Esta ontología (Fig. 1) se podrá aplicar para seleccionar las preguntas del cuestionario en su presentación al cliente o para el almacenamiento de las preguntas y respuestas en una base de datos.

Aunque incluye conceptos aplicables a cualquier sistema de encuestas, podría ser necesario ampliar los términos definidos para utilizar esta ontología en otros campos estadísticos.

4.2 Ontología de Calidad de Servicio

Esta ontología (Fig. 2) debería servir como base para la automatización de procesos que utilicen conceptos relacionados con la calidad de servicio. En ella se localizan los siguientes términos:

- Servicio. Un servicio que se proporciona. Podría no definirse en esta ontología y utilizar una de las ontologías de servicios disponibles en la web como OWL-S [29].

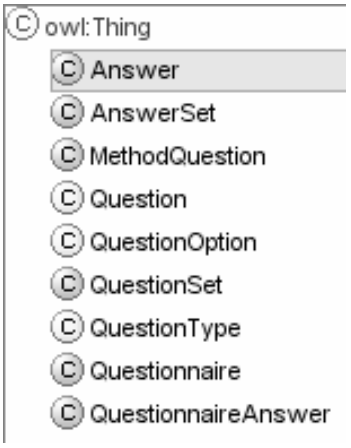


Figura 1. Vista de las clases de la ontología de cuestionarios en Protégé [28]

- Proveedor de servicio. Empresa o particular que proporciona un servicio. Podría no definirse en esta ontología y utilizar una de las ontologías de servicios disponibles en la web como OWL-S [29].

- Calidad de servicio. Concepto de calidad de un servicio.

- Atributo de calidad de un servicio. Parámetro que se utiliza para calcular la calidad de un servicio y que será medido.

- Valor de la calidad de un servicio. Cálculo de la calidad de un servicio en una fecha determinada a partir de los atributos medidos.

- Valor de atributo de la calidad de servicio. Medida de uno de los atributos o parámetros objetivos o subjetivos que sirven para determinar el valor de la calidad en una fecha determinada.

- Tipo de medida. Si esta se realiza con técnicas basadas en parámetros objetivos o subjetivos. Existen dos instancias: objetiva y subjetiva.

- Método de medida. Técnica que se utiliza para medir la calidad de un servicio o los parámetros de la misma. Incluye una instancia correspondiente al SLA de un proveedor.

4.3 Ontología ServQual

Los conceptos incluidos en esta ontología (Fig. 3) dan soporte al método ServQual de medición de la calidad de servicio percibida. Se han separado de las otras dos ontologías por su carácter específico, pero se basará en ellas.

Los términos principales que define son los siguientes:

- Atributo ServQual, que extiende a los atributos de calidad de servicio restringiendo a los cinco de ServQual.

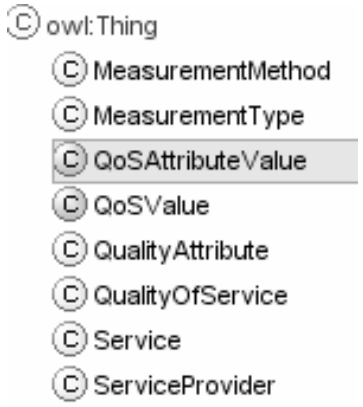


Figura 2. Vista de las clases de la ontología de calidad de servicio en Protégé [28]

- Conjunto ServQual de respuestas y de preguntas. Extienden los grupos de preguntas y respuestas de los cuestionarios incluyendo una propiedad que relaciona de cada grupo con un atributo ServQual.

- Cuestionario de un servicio. El instrumento ServQual que ha de estar compuesto por conjuntos de preguntas ServQual.

- Dimensión de usuario. Define las dimensiones de percepción o expectativa para poder diferencias grupos de preguntas.

- Valor medio de un atributo ServQual. Valor de un atributo para un formulario obtenido según el método ServQual, es decir, haciendo media entre las preguntas relacionadas con ese atributo.

- Media ServQual. Valor ServQual para un formulario obtenido como media ponderada de los atributos.

Además incluye una instancia de método de medida de calidad llamada ServQual correspondiente al

Figura 3. Vista de las clases de la ontología ServQual en Protégé [28]

cálculo la calidad de servicio usando este método basado en parámetros subjetivos.

El uso de esta ontología combinada con las dos vistas previamente facilitará la automatización de la captura de datos de opinión de usuario para el cálculo de la calidad de servicio percibida.

5 Comparativa de Ontologías

Existen diversas propuestas relacionadas con la creación de ontologías para dar soporte a la calidad de los servicios. A continuación se realiza una comparativa entre las ontologías desarrolladas y las existentes.

Es necesario comentar, que ninguna de las ontologías que se van a describir centra su enfoque en la calidad de servicio percibida, y que la comparación podrá establecerse sólo con la ontología de calidad de servicio de este trabajo (dejando a un lado la ontología de cuestionarios y la ontología ServQual).

[21] realiza una aproximación orientada a servicios web en los que el usuario pueda solicitar una calidad y se establezca una comparación con la calidad anunciada por el proveedor. La calidad de servicio anunciada se basa en propiedades que pueden ser de entrada (por ejemplo: input bit rate), de salida (output bit rate) o generales (core) dependiendo de su naturaleza, además de precondiciones que han de darse para que se cumpla el nivel prometido y los efectos esperados. Por último, ofrece la definición del concepto de métrica para expresar cómo se han de realizar las mediciones que comprueben que se cumple la calidad de servicio anunciada.

La ontología definida está más orientada a expresar información por parte del proveedor a través de un documento (un acuerdo de nivel de servicio SLA, por ejemplo), frente al aquí presentado que se encuentra más enfocado a la medición de parámetros para calcular la calidad en un momento dado.

[14] considera una ontología bastante similar a la presentada en este artículo, si bien no define el concepto de método de medida para obtener una medida concreta. Por otro lado, la forma en que luego la desarrolla para un caso concreto (web services) es diferente a la que se aplica aquí para ServQual, sobre todo en cuanto a la definición de algunos términos como clases o como instancias.

SUMO (Suggested Upper Merged Ontology) y MILO (MId-Level Ontology) [15] son ontologías que incluyen algunos términos útiles como el de unidad de medida, pero que no entran en detalle sobre la calidad de servicio. La ontología de servicios informáticos (*Ontology of Services*) basada en ellas proporciona algún concepto adicional como los actores (sensores y monitores) y algunas características (ancho de banda, por ejemplo), pero tampoco engloba los elementos aquí presentados.

Por otro lado, existen un conjunto de ontologías que han tratado de definir la calidad de servicio para el servicio de red. Destacan la ontología de FIPA (Foundation for Intelligent Physical Agents) [7] y OMNI/IP (Ontology for Metrics of Network Infrastructure using the Internet Protocol) [8].

FIPA permite a los agentes inteligentes intercambiar información de QoS. Se utilizan términos bien conocidos en redes como latencia, variación del retardo (*jitter*) o tasa de error (*BER*). Además define tres posibles tipos de valores para los parámetros que pueden medirse en tiempo (horas, minutos, segundos o milisegundos), probabilidad o tasa de datos (gbits/s, mbits/s, kbits/s y bits/s). Además define términos para describir el canal de comunicación y el protocolo que se usa. Una adaptación de esta ontología se usa en [18].

OMNI/IP busca la compatibilidad con el estándar IETF/IPPM (IP Performance Metrics) [10]. En este artículo se expone que la ontología de FIPA está menos centrada en los problemas de red que OMNI/IP y que sólo unos pocos de sus términos están incluidos en estándares aceptados. La representación UML de la ontología indica que se presentan conceptos de red como *Interface* o *Host* junto a una clase llamada *PerformanceMetric* de la que heredan las métricas de *Connectivity*, *RoundTripDelay*, *OneWayDelay*, *DelayVariation* y *OneWayPacketLoss*. En los ejemplos de utilización se observa que las propiedades incluidas en la ontología para estas métricas están relacionadas con las de los estándares mencionados. Sin embargo, la falta de disponibilidad de una versión de esta ontología en las páginas del proyecto impide su correcta validación.

Ambas ontologías, como se ve, se centran en un servicio concreto, frente a la aquí presentada. Podría hacerse una integración de ellas con la de calidad de servicio descrita en este texto. Para ello, en la de FIPA se interpretarían los parámetros de red como subclases o instancias del atributo de calidad de servicio. En OMNI/IP, *PerformanceMetric* podría fundirse con el método de medida de la ontología de calidad de servicio o ser una subclase de ella. Además se podría conseguir expresar las relaciones entre los parámetros subjetivos de la ontología ServQual presentada en este artículo y las ontologías específicas de servicio FIPA y OMNI/IP mediante el uso de funciones de utilidad. Estas funciones de utilidad tendrían que expresarse mediante alguna extensión a los lenguajes de ontologías, como Semantic Web Rule Language (SWRL) [27] que permita incluir reglas.

6 Intercambio de datos de calidad de servicio percibida

Las ontologías presentadas en este artículo han de servir para automatizar la obtención de la calidad de

servicio percibida por los usuarios. Para ello se definen un conjunto de arquitecturas que podrían implementarse y que utilizarían estas ontologías para intercambiar los datos o almacenarlos:

- Interacción cliente-proveedor: Los proveedores (o un tercero independiente) publican los cuestionarios ServQual. Una aplicación residente en los ordenadores de los clientes (o en un servidor de encuestas) consulta los servidores localizados en los proveedores de servicios para obtener esos cuestionarios. Cuando los clientes contestan a los cuestionarios, se produce una transferencia de información en sentido contrario entre los ordenadores de los clientes (o servidores de encuestas) y los proveedores. Se pueden solicitar las respuestas completas o el valor ServQual ya calculado. El proveedor almacena los datos en su repositorio.

- Cooperación entre clientes: Similar a la anterior, en este caso los clientes comparten sus respuestas mediante una aplicación peer-to-peer. De esta manera los usuarios pueden decidir qué proveedor es el más conveniente para un servicio determinado sin necesidad de hacer uso de un sistema de intermediación.

La primera arquitectura proporciona realimentación a los proveedores. La segunda permite comparar a los clientes los servicios proporcionados por su proveedor con los de la competencia.

La implementación podría basarse en una arquitectura de agentes inteligentes utilizando la tecnología Nuin [25] basada en el entorno de agentes Jade (Java Agent DEvelopment framework) [26] y el entorno para la web semántica Jena [23], aunque todavía está en una versión muy básica. En la actualidad se está realizando una implementación con Joseki [24] y Jena[23].

7 Conclusiones

Existe una necesidad de capturar la opinión de los usuarios acerca de un servicio para disponer de la calidad de servicio percibida, puesto que esta difiere de la calidad proporcionada.

En este artículo se ha presentado un conjunto de ontologías que permiten automatizar la captura de datos de calidad percibida basándose en la metodología ServQual. La ontología de calidad de servicio definida profundiza en los aspectos relacionados con las medidas de la calidad, aspecto poco tratado en otros trabajos existentes. Por otro lado, las ontologías de cuestionarios y ServQual definidas son las que permiten el tratamiento de la información sobre la calidad percibida del servicio.

De este modo, se dispone de las herramientas básicas que permitirán profundizar en la relación existente entre calidad percibida y parámetros objetivos del servicio. El siguiente paso consiste en el diseño de experimentos que permitan obtener una muestra significativa de información sobre calidad percibida y los parámetros objetivos que dan lugar a esa calidad percibida, con el objetivo de encontrar las funciones de utilidad que representen la relación entre los dos aspectos.

Para ello, se han sugerido dos arquitecturas alternativas para la aplicación de las ontologías en un entorno de provisión de servicios basados en red. La implementación de una de estas arquitecturas basándose en herramientas de web semántica o agentes inteligentes ya disponibles parece factible. Sin embargo, la selección de una de ellas para su integración y posterior experimentación requiere de un análisis más exhaustivo, que contemple no sólo los aspectos tecnológicos, sino también los económicos y la infraestructura disponible.

Agradecimientos

Este trabajo ha sido financiado en parte por el Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica (I+D+I) y el Ministerio de Ciencia y Tecnología, a través del proyecto TIC2003-04406.

Referencias

[1] Bahtti, N., Bouch, A., Kuchinsky, A. "Integrating User-Perceived Quality into Web Server Design". Hewlett Packard Laboratories. Palo Alto, EEUU, Enero 2000. http://www.hpl.hp.com/techreports/2000/HPL-2000-3.pdf

[2] Bahtti, N., Bouch, A., Kuchinsky, A. "Quality is in the Eye of the Beholder: Meeting User's Requirements for Internet Quality of Service". Hewlett Packard Laboratories. Palo Alto, EEUU, Enero 2000. http://www.hpl.hp.com/techreports/2000/HPL-2000-4.pdf

[3] Bouch, A., De Meer, H., Sasse, A. "Of Packets and People: A User-centered Approach to Quality of Service". Proceedings of International Workshop on QoS. 2000. http://www.cs.ucl.ac.uk/staff/A.Sasse/iwqos2000.ps

[4] Bouch, A., Sasse, A. "It ain't what you charge, it's the way that you do it: A user perspective of network QoS and pricing". Proceedings of the IFIP/IEEE International Symposium on Integrated Network Management (IM'99), pp. 639-655. Boston, EEUU, Mayo 1999.

[5] Pastor, E., Sánchez-Macián, A. "Evaluación de la Calidad de Servicio en Redes IP. Percepción de Usuario y Métricas". Actas de las XIV Jornadas Telecom I+D. Madrid 2004

[6] Cody, K., Hope, B. "EX-SERVQUAL. An instrument to measure quality if Extranets." Proceedings 10th Australasian Conference on Information Systems. 1999. http://www.vuw.ac.nz/acis99/Papers/PaperCody-169.pdf

[7] Foundation for Intelligent Physical Agents. *FIPA Quality of Service Ontology Specification*. Noviembre 2002 http://www.fipa.org/specs/fipa00094/XC00094.html

[8] Herfurt, M. "OMNI/IP – Ontology for Metrics of Network Infrastructure using the Internet Protocol". http://agentsmith.salzburgresearch.at/Downloads/NetworkOntology.pdf

[9] Intelligence, Agents, Multimedia (IAM) Group. *Currency Ontology*. University of Southampton. Marzo 2003. http://www.daml.ecs.soton.ac.uk/ont/currency.owl

[10] IETF Network Working Group, RFC 2330 – Framework for IP Performance Metrics. Mayo 1998.

[11] Iwaarden, J. van, Wiele, T. van der. "A study of the applicability of SERVQUAL dimensions for web sites". Erasmus University. Rotterdam, Holanda. 2003. http://www.few.eur.nl/few/people/vaniwaarden/publications/0007.pdf

[12] Knowledge Systems, AI Laboratory (KSL) *Ontolingua Server*. Stanford University. http://www-ksl-svc.stanford.edu:5915/

[13] Ludwig, H. "Web Services QoS: External SLAs and Internal Policies Or: How do we deliver what we promise?" Fourth International Conference on Web Information Systems Engineering Workshops (WISEW'03) Diciembre 2003. http://alarcos.inf-cr.uclm.es/wqw2003/Abstract%20Ludwig.pdf

[14] Maximilien, M., Singh, M. "A framework and ontology for Dynamic Web Services Selection". IEEE Internet Computing. Septiembre 2004. http://www.csc.ncsu.edu/faculty/mpsingh/papers/mas/IEEE-IC-selection-sep-04.pdf

[15] Niles, I., and Pease, A. 2001. "Towards a Standard Upper Ontology". Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001), Ogunquit, Maine. Octubre, 2001. http://ontology.teknowledge.com/

[16] Saaksjarvi, M., Saarinen, T. "Evaluation of Service Quality of Information Systems." Proceedings of the Second International Software Metrics Symposium, pág 84-94. IEEE. 1994

[17] Sumra R., Arulazi D. "Quality of Service for Web Services—Demystification, Limitations, and Best Practices". Marzo 2003 http://www.developer.com/services/article.php/2027911

[18] Toivonen, S., Helin, H., Laukkanen, M., Pitkäranta, T. "Context-Sensitive Conversation Patterns for Agents in Wireless Environments". Proceedings of the International Workshop on Ubiquitous Computing (IWUC'04). Porto, Portugal. Abril 2004, pp. 11--17. http://www.vtt.fi/tte/staff/tos/publications/ubiip.pdf

[19] Tosic V, Esfandiari B., Pagurek B., Patel K. "On Requirements for Ontologies in Management of Web Services". 2002. http://www.csd.uch.gr/~hy565/Papers/requirements_for_ontologies.pdf

[20] Villalobos, G. "Web-Application for the Customer Satisfaction Measurement". Thesis. Fribourg University. Swiss. Octubre 2000: http://www.inf.ufrgs.br/granvile/QoS/Imprimir/faq.htm

[21] Zhou, C., Chia, L-T., Lee, B-S. "DAML-QoS Ontology for Web Services". Proceedings of the IEEE International Conference on Web Services (ICWS'04), Junio 2004. http://www.ntu.edu.sg/home5/PG04878518/Articles/icws04_235_Chen_Z.pdf

[22] Zeithaml, V. A., Parasuraman, A., Berry, L. L. (1990). "Delivering Quality Service. Balancing Customer Perceptions and Expectations". The Free Press. Macmillan, Inc. New York. United States.

[23] Jena – A Semantic Web Framework for Java. http://jena.sourceforge.net/

[24] Joseki – The Jena RDF Server. http://www.joseki.org/

[25] Nuin – The Jena Agent Framework. http://www.nuin.org/

[26] Jade – Java Agent DEvelopment Framework. http://jade.tilab.com/

[27] I. Horrocks, P. F. Patel-Schneider, H. Boley, S. Tabet, B. Grosof, and M. Dean, "SWRL: A Semantic Web Rule Language Combining OWL and RuleML". http://www.w3.org/Submission/SWRL/

[28] The Protégé Ontology Editor and Knowledge Acquisition System.

[29] OWL-S. A OWL-based Web service ontology. http://www.daml.org/services/owl-s/1.0/

Hacia una plataforma semántica para servicios de eGovernment

Luis A. Sabucedo y Luis Anido
Departamento de Ingeniería Telemática. Universidad de Vigo
ETSI de Telecomunicación. C/ Maxwell S/N. Campus Universitario.
36310 - Vigo (Pontevedra)
Teléfono: 986 81 40 73 Fax: 986 81 21 16
E-mail: {lsabucedo,lanido}@det.uvigo.es

Abstract *Currently, we are witnessing a soaring process of eGovernment projects in the so-called developed world. This lead us to a situation where services are no compatible among them and they fulfill limited features as they are intended to provide short term solutions. This paper proposes a semantic-based platform for the provision of facilities oriented towards holistic, transactional and interoperable services where citizens are the center of the processes in their relations with Public Administrations.*

1. Introducción

La revolución tecnológica que representa la aplicación de las TICs (Tecnologías de la Información y Comunicación) a los procesos habituales en las diferentes interacciones humanas ha recorrido una serie de etapas hasta llegar al momento actual. De entre los primeros campos de aplicación de las TIC a procesos productivos humanos podemos significar el eBusiness, que permitía aplicar procedimientos electrónicos a la gestión empresarial, de la mano del uso de EDI. Conforme el coste de los equipos electrónicos fue disminuyendo y su potencia de operación creciendo de un modo exponencial el ámbito de las TIC fue aumentando y colándose en más campos de aplicación. De este modo, hemos sido testigos de la aparición de más *eTechnologies* tales como eCommerce, eLearning, eHealth, ... Conforme la madurez de estas tecnologías ha ido evolucionando de la mano de la experiencia y el progreso se ha tendido paulatinamente a hacer un uso más extensivo de las TICs. El ejemplo más extremo de esto es la aparición del eGovernment.

Al hablar de eGovernment no nos estamos refiriendo simplemente a la prestación de servicios administrativos del Estado mediando interfaces del entorno de las TICs. En realidad, el eGovernment tiene como objetivo un complejo proceso de reingeniería que pretende realizar una reorientación de los servicios que prestan los gobiernos para situar en su centro a los ciudadanos (los clientes más habituales de estos sistemas) en vez de considerar los procesos administrativos cómo un requisito organizativo.

Dado el potencial de estas tecnologías, la mayor parte de los estados lleva invirtiendo recursos estos últimos años. Fruto de esta cantidad de esfuerzo dedicado a este campo de investigación y desarrollo, nos podemos encontrar gran cantidad de soluciones y sistemas que proporcionan soluciones reales. De este modo, los países del llamado primer mundo prestan servicios telemáticos para el entorno del eGovernment, podemos encontrar una revisión rigurosa del estado en que se encuentra cada país en revisiones exhaustivas de organismos internacionales como la ONU[11] que nos permiten tener una visión de la situación de cada país de una manera global y comparativa.

El problema que nos encontramos es la falta de interoperabilidad y capacidad para la extensión de servicios de un modo sistemático y organizado. Dado que cada país presta sus servicios sin considerar la situación que se da en los países de su entorno, nos encontramos con modelos de datos diferentes en cada caso y una división funcional de servicios que tampoco suele estar realizada de un modo homogéneo. De hecho, nos podemos encontrar soluciones desarrolladas en el marco de la misma Administración Pública, de aquí en adelante AP, que no son interoperables.

Para hacer posible el desarrollo de plataformas de servicios integradas en las que se puedan desarrollar proyectos de alto valor añadido se hace necesario proporcionar un soporte que pueda integrar estos sistemas. En este caso y dado que estos proyectos se enfocan desde un punto de vista de modelos de datos y funciones disponibles a nivel de recurso software, nuestra propuesta se centra en la provisión de soluciones desde un nivel superior del sistema. Nuestro objetivo es la provisión de estos servicios en una capa semántica.

El presente artículo enfoca el problema de diseñar y desarrollar una infraestructura donde prestar servicios de eGovernment orientada hacia agentes móviles e inteligentes que puedan interactuar para la prestación de servicios empleando para ello los recursos que nos facilita el uso de una infraestructura semántica para referenciar y ejecutar recursos que permitan mejorar el nivel de penetración de los servicios de eGovernment dentro de la sociedad.

Este artículo se organiza de la siguiente manera: en primer lugar veremos una introducción al concepto de información semántica frente a modelos basados en datos; luego veremos un repaso al estado del arte de los proyectos de eGovernment. Después haremos nuestra propuesta de Arquitectura de servicios. Esta propuesta se organiza en una introducción a la propuesta de solución; una revisión de los agentes involucrados en el sistema, una primera aproximación al desarrollo del sistema y un acercamiento a una implementación del sistema. Finalmente, se extraen conclusiones obtenidas sobre el modelo propuesto.

2. Abstracción de implementaciones: la web semántica

El problema al que tenemos que hacer frente es la provisión de servicios en un entorno muy particular en el que hay dificultades dado lo heterogéneo de las funcionalidades ya propuestas y los requisitos particulares que se presentan en cuanto a accesibilidad y seguridad de los servicios que hay que prestar al cliente del sistema, el ciudadano. Para hacer esto posible la mejor opción que tenemos a nuestro alcance es el uso de la Web Semántica.

La Web Semántica es un concepto introducido por Tim Berners-Lee en el 2001[3]. La idea que subyace en esta definición es hacer una transición desde la provisión de contenidos basados en datos hacia el soporte de información. Es decir, pasar de datos que puedan ser intercambiados y procesados por los humanos usuarios hacia información procesable por las propias máquinas. Este concepto permite la automatización y mecanizado de gran cantidad de operaciones que proporcionan servicios de valor añadido de un modo completamente mecánico por parte de entidades *ordenador* dotadas de algún tipo de lógica programada. Es decir, si bien un modulo software no puede hacer inferencias sobre documentos expresados en XML cuando los tags no son exactamente iguales sí es posible que procese documentos si en ellos se expresan conocimientos respaldados por sistemas basados en conocimiento.

Usando una aproximación al problema desde el punto de vista semántico podemos conseguir abstraer la problemática a un plano superior desde el que hacer una planificación interoperable y donde tienen cabida procesos automatizados de alto valor añadido sin intervención humana, o al menos, siendo ésta mínima.

Una de las piezas clave de los sistemas basados en conocimiento es el uso de Ontologías. Estas nos permiten la fijación de conceptos del entendimiento humano a soportes informáticos para su posterior tratamiento de un modo automatizado. Una definición para el concepto de ontología muy ajustada y de gran aceptación en nuestro entorno es la formulada en [8]:

> "An ontology is a formal, explicit specification of a shared conceptualisation of a domain of interest."

En esta definición encontramos reflejadas algunas de las características más relevantes que se encuentran en una ontología. De este modo, podemos destacar que una ontología es un modo formal (ateniéndose a una formulación preestablecida y rigurosa) de plasmar una conceptualización (conocimiento humano sobre un tema) dentro de un ámbito de conocimiento específico. De esta definición ya podemos observar el primer, y quizás más grande, problema que hay que afrontar a la hora de establecer ontologías en un dominio: se trata de juntar el conocimiento del dominio (requiere personas del ámbito del problema) y conocimiento propio de la Web Semántica para poder plasmar los conocimientos de una manera formal. Para plasmar esta concretización de los conocimientos del dominio se ha evolucionado a través de diferentes lenguajes:

RDF and RDF Schemas (Resource Description Framework) [6]. Un formato del W3C[1] basado en XML que permite la descripción de contenidos basándose en metadatos.

DAML (DARPA Agent Markup Language)[18]. La podemos considerar como una extensión de RDF para expresar de un modo semántico información sobre recursos.

OIL (Ontology Interchange Language and Ontology Inference Layer)[13]. Es el resultado del proyecto On-To-Knowledge[2] enmarcado en las iniciativas de European IST. Está orientado hacia la presentación de contenidos dentro del entorno de los recursos Web. Se considera superado por el uso de DAML+OIL.

OWL (Ontology Web Language)[20]. Es una recomendación del W3C que cubre todos los aspectos presentes en DAML+OIL y está orientado a proporcionar un modo altamente funcional de expresar ontologías. Además aporta un alto grado de flexibilidad al permitir la articulación de la información en diferentes niveles: OWL Lite, OWL DL, OWL Full. Estos diferentes niveles permiten un acceso diferenciado a la información en función de las necesidades de cada caso.

Usando cualquiera de estas tecnologías podemos describir conocimientos semánticos con un soporte formal bastante robusto y ayudados por

[1] http://www.w3.org
[2] http://www.ontoknowledge.org/

herramientas bastante sólidas. No obstante, hay un problema adicional que también deberemos considerar desde la óptica de la web semántica: la descripción de servicios. Si bien estas ontologías nos permiten la definición de conocimiento genérico hay una funcionalidad básica en estos sistemas que debe ser cubierta de un modo similar. Una necesidad habitual en estos entornos es el descubrimiento e invocación de servicios que también debería ser posible con este tipo de soporte. Para ello, hay disponibles varias opciones. Las más destacadas son:

OWL-S (Ontology Web Lenguage - Services)[5]. Se trata de una especificación de OWL especializada en la descripción de servicios Web. Mediante esta tecnología la información se organiza en tres niveles para la descripción de los diferentes aspectos del servicio descrito y facilitar su recuperación y orquestración.

WSMO (Web Service Modeling Object)[14]. Esta propuesta es el fruto del trabajo del grupo de trabajo SDK WSMO. El objetivo de esta tecnología es la provisión de un modo formal de describir los servicios presentes en el sistema de un modo semántico.

En la selección de una u otra opción hay que considerar varios factores de un modo conjunto. En primer lugar hay que considerar la potencia semántica de cada opción, pero también la cantidad de software disponible para cada una de ellas y la integración con el resto del sistema. Para nuestra propuesta nos hemos decantada por el uso de OWL-S dado que nos permite reutilizar conceptos definidos mediante OWL y disponer de software con un nivel de funcionalidad bastante alto como por ejemplo Jena[9] o Protégé[16] y plugins relacionados[17].

3. Entorno de la solución

Como ya se ha indicado anteriormente, a día de hoy existe gran cantidad de proyectos que pretenden cubrir demandas de los ciudadanos en lo que a relación con la administración se refiere, soluciones del ámbito de G2C (Government To Citizen). Estas soluciones podemos clasificarlas esencialmente:

Soluciones basadas en *front-end*. Este tipo de soluciones proporcionan servicios puntales pero acceden a registros de transacciones o ciudadanos por lo que no son adecuados para soluciones de gran tamaño.

Soluciones basadas en *back-end*. Estas soluciones trazan las operaciones y acceden a registros de los usuarios pero no permiten un acceso a ellos a los ciudadanos por lo que se hace necesario la intervención de los empleados de la administración pública.

Este enfoque para el planteamiento de soluciones no es el más adecuado ya que nos conducen a soluciones de un horizonte temporal corto y con una baja interoperabilidad. Existen también gran cantidad de soluciones y marcos para el desarrollo de productos dentro de este ámbito en las diferentes Administraciones Públicas (SAGA[10], FEAF[15],eGIF[19],ADEA[7]). La mayor parte de estas propuestas se basan en realizar análisis y proponer soluciones basadas en datos ya que su objetivo es el proporcionar un entorno en el que desarrollar soluciones a día de hoy, y dado el estado actual de los modelos semánticos para proyectos en el corto plazo es el mejor enfoque.

Sin embargo, teniendo en cuenta las actuales líneas de investigación, es claro que se dedicaran recursos a la implementación de sistemas basados en conocimiento para la provisión de soluciones en el área del eGovernment, dando lugar al ya conocido como kGovernment[3].

Dentro de esta línea de proyectos podemos destacar:

OntoGov[4]. El objetivo de este proyecto es la provisión de una ontología que refleje el conocimiento necesario para su aplicación al eGovernment.

Terregov[5]. Este proyecto europeo tiene como finalidad la provisión de una interfaz web homogénea para diferentes servicios de eHealth mediante el uso de semántica aplicada al dominio.

eGoia[6]. Se trata de un proyecto marco que pretende la prestación servicios de eGovernment en el largo plazo orientado hacia los ciudadanos de una manera sencilla y barata.

4. Solución propuesta

Las soluciones desarrolladas para el entorno del eGovernment tienen sus propios requisitos que si bien no son nuevos frente a otros entornos sí que tienen unas condiciones particulares que deben ser satisfechas:

- Se hace necesario prestar servicios de no repudiación tanto en origen como en destino. Esto es importante ya que los acuses de recibos firmados que debemos facilitar tiene

[3] Especilización dentro del área del eGovernment que pretende el uso de sistemas inteligentes basados en conocimiento para la provisión de soluciones.

[4] `http://www.ontogov.com/`

[5] `http://www.terregov.eupm.net/`

[6] `http://www.egoia.info/`

valor jurídico y pueden ser requeridos como modo de acceso a posteriores recursos.

- El sistema debe permitir una completa trazabilidad de las operaciones desarrolladas, así como debe facilitar mecanismos para auditar estas operaciones.
- Los clientes del sistema pueden exigir el acceso al código fuente de los clientes que ejecutan dada la necesidad extrema de confiabilidad del sistema.
- Se debe proporcionar un modo de acceso a todos clientes sin imponer restricciones en función de plataformas software o hardware en la medida de los posible. Incluso hay que preveer el caso de usuarios con limitaciones sensoriales, y proporcionar contenidos o soporte adaptado a estos.

Para hacer esto posible, en nuestra propuesta señalamos como características deseables:

- Open Source. Nuestro sistema debe proporcionar soluciones abiertas como medio de asegurar el correcto trato de la información gestionada. Esto permitirá ganar la confianza de los usuarios que podrán ver que la solución está implementada correcta y favorece que haya nuevas aportaciones al sistema.
- Adopción de estándares abiertos. Dado que debemos facilitar soporte para acceso universal no podemos ligar las soluciones a desarrollos propietarios y que pueden obligar a la compra de productos por parte de los clientes. Además, mediante el uso de estándares abiertos en todas las capas (software, comunicaciones, almacenamiento, ...) estamos en condiciones de garantizar la contuidad del proyecto sin necesidad de atarse a soluciones tecnológicas puntuales.
- Uso de agentes independientes para construir la red. Las operaciones serán llevadas a cabo por agentes independientes en el sistema que trabajaran de un modo coordinado para alcanzar la solución del problema en cada caso.
- Uso de Conocimiento para gestionar el sistema. Dado lo heterogeneo del sistema el mejor modo de tener definidos los servicios y los agentes es mediante el uso de una ontología en la que queden registrados todos las posibilidades del sistema y permita a los nuevos agentes en el sistema acceder a la información y recursos del sistema.

Estas características de diseño nos permite disponer de una red de agentes que gestionaran por ellos mismos los servicios que se le demanden. De este modo podemos hablar de una solución para mGovernment[7]. Es decir, los agentes que integran nuestra red serán independientes y en ellos podemos introducir toda la lógica que deseemos. De hecho, en el *workflow* general del sistema los agentes cliente del sistema tendrán asignadas gran cantidad de las responsabilidades incluyendo las tareas de descubrimiento y coreografía de las funciones que necesite invocar

5. Agentes en el sistema

Según lo expuesto hasta ahora podemos establecer que nuestro sistema dispondrá de diferentes tipos de agentes (ver Fig. 1).

5.1. Agentes cliente

Los elementos más relevante es el agente cliente en tanto en cuanto es el responsable de invocar los servicios y es el que debe recibir los resultados de las operaciones que se ejecutan en los proveedores de servicios. Estos agentes deberán ser responsables de descubrir los servicios existentes en la red (ver Servidor de Páginas Azules) coordinar las llamadas a los Proveedores de servicios y completar los servicios.

De este modo, dispondremos de un componente software altamente independiente capaz de organizar el *workflow* de las operaciones de un modo autónomo. Las ventajas de esta característica son que nos permite desarrollar agentes dotados de toda la funcionalidad que deseemos y bajo cualquier entorno. Debemos pensar que esta solución que planteamos valdrá para el desarrollo de agentes que se ejecuten tanto bajo la forma de aplicación local en un ordenador de sobremesa como para el desarrollo de aplicaciones MHP que se ejecuten en un decodificador de TV dentro de la plataforma.

Tal y como está definido el sistema para la prestación de un servicios basados en una interfaz Web (o WAP) deberemos desarrollar un agente que realiza las funciones de un front-end y genere la interfaz para la presentación del HTML (o WML) mediante un comportamiento tipo proxy.

5.2. Servidor de páginas azules

Este elemento de la arquitectura de servicio es el encargado de facilitar las búsquedas de los servicios en el sistema. Para ello, este agente está dotado de un repositorio de contenidos en el que se define semánticamente los servicios que presta. De este modo los agentes se dirigen a él para localizar los servicios y/o servidores que desean invocar enviando una descripción semántica de qué es lo que desean localizar. Este sistema se encarga de

[7] Especilización dentro del área del eGovernment que pretende el desarrollo de soluciones aplicables a entornos móviles, típicamente teléfonos móviles, PDAs, plataformas MHP (*Multimedia Home Platform*), ...

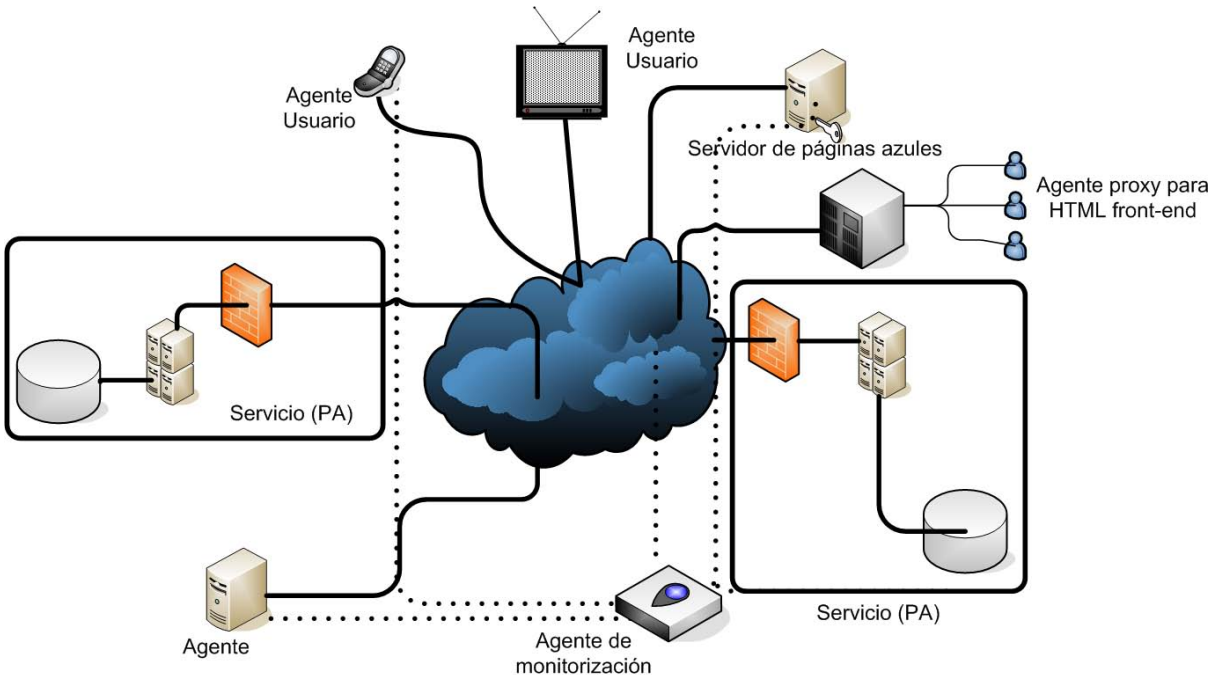


Figura 1: Agentes en el sistema

localizar y devolver aquellos que encajen en su selección, no siendo de su responsabilidad como se orquesten las llamadas en el caso de servicios no atómicos.

5.3. *Brokers*

Igual que en otros muchos entornos donde proliferan los servidores de servicios relacionados, da lugar la aparición de estos componentes que tiene como finalidad la provisión de servicios de alto valor añadido o incluso proporcionar interoperabilidad frente a sistemas no propios del sistema. De este modo, se hace posible que se presenten servicios como la búsqueda avanzada de recursos (ofertas de empleo, subvenciones, becas, ...) o el acceso a operaciones automatizadas frente a diferentes APs (recuperación de toda la información personal en diferentes APs, procesos de alta o baja ante diferentes APs, ...).

Obviamente, y del mismo modo que en otros dominios donde se emplean *brokers*, se hace necesario el uso de una interfaz propia para los *brokers* que permita la expresión las solicitudes que puede recibir y la manipulación de los diferentes modelos de datos que gestiona el sistema.

5.4. Proveedores de Servicio

Estos elementos del sistema tienen como responsabilidad la prestación final de los servicios y su implementación y gestión es responsabilidad de las APs correspondientes. De este modo, podemos delegar en ellas la responsabilidad de la prestación de estos servicios a través de una interfaz que permita la invocación de los servicios provistos. Por supuesto, esta interfaz y los servicios mismos que presta deben encajar en la ontología que define el funcionamiento global del sistema.

5.4.1. Capa de convergencia

Se hace necesario introducir una capa intermedia que realice las tareas necesarias para la convergencia entre la capa semántica del sistema y la capa basada en datos. Dado que la solución propuesta resuelve el problema subiendo el nivel en el que trabajamos hasta la capa semántica, se hace necesario que en algún punto del sistema se produzca la vuelta hacia el nivel en el que los servicios son realmente facilitados por los sistemas que nos podemos encontrar en realidad. Esta funcionalidad estará típicamente relacionada con los servidores por ser los elementos ya existentes y que operan en el nivel de servicio invocado o intercambio de datos.

6. Desarrollo del sistema

Para el diseño de un sistema de las características propias que se encuentran en este entorno (muy alta escala, entorno altamente dinámico, riesgos de seguridad, ...) debemos asumir algunos problemas: ausencia de una única herramienta para el diseño, aproximaciones mediante casos de uso pueden ser problemáticas, el uso de XML como formato de datos no siempre es el óptimo, existe la necesidad de proporcionar una documentación en un formato aceptado por todos los agentes involucrados, ...

El modelo que se plantea para el diseño de alto nivel se hará siguiendo el modelo SOA (*Servi-*

ce Oriented Architecture). Haciendo uso de PUM (Proceso Unificado de Modelado)[4] y a las aportaciones hechas por Bass et al[2] estaremos en condiciones de seguir una metodología precisa que nos permitirá abordar el problema con garantías de éxito. Si bien no existe ninguna restricción a priori a resultas de estos modelos de desarrollo, se ha seleccionado una implementación basada en Web Services para el desarrollo del sistema; para esto se siguen las recomendaciones del W3C para estas arquitecturas[21].

7. Acercamiento a una implementación

La implementación de un sistema tan complejo como esto es bastante complicada y debe plantearse considerando no sólo soluciones tecnológicas puntuales si no también la problemática que puede surgir en futuro a resultas del cambio de tecnología, por problemas de escalabilidad o modificación de requisitos de diseño.

En nuestro caso particular hemos optado por las siguientes opciones y metodologías para las diferentes partes del sistema.

7.1. Ontología

El único nexo de unión entre todos los agentes en el sistema, hablando a nivel de aplicación, es la ontología. Esta ontología almacena toda la información necesaria para hacer que cualquier agente en el sistema sea capaz de expresar sus necesidades y capacidades de modo tal que los demás componentes de la red puedan entenderlo y colaborar con él en el cumplimiento de su cometido.

En nuestro caso para la implementación del sistema hemos recurrido a la explicitación de los conocimientos del sistema mediante el uso del lenguaje OWL. Este nos permite expresar de un modo bastante completo los diferentes eventos y agentes en el sistema.

La definición de una ontología completa que permita modelar de un modo aproximado el comportamiento del sistema es una tarea complicada que está sujeta a procesos de cambio de muy rápidos tanto por motivos técnicos (modificaciones de la funcionalidad disponible) como por motivos legales/políticos (cambios en los servicios prestados). Se hace necesario incluso considerar cuidadosamente los casos en los que no sea posible dar una respuesta apropiada e informar adecuadamente a los usuarios de la situación a fin de evitar situaciones de indefensión en los ciudadanos.

Por ese motivo y la complejidad de la información a tratar resulta muy costoso su desarrollo. Para conseguir esta ontología la metodología propuesta está basada en un acercamiento partiendo de los eventos que pueden tener lugar en la vida de un ciudadano y a partir de estos se sintetiza todo el modelo de conocimiento.

A la hora de decantarse por una opción tecnológica para la definición de una ontología hay que valorar diferentes aspectos: mantenibilidad de la información, herramientas disponibles para la gestión, posibilidades de uso, escalabilidad del sistema. Por esos motivos hemos optado por el uso de OWL ya que nos permite el modelado del sistema de un modo razonablemente sencillo y tenemos disponibles herramientas para la gestión de la información. De este modos hemos establecido una ontología en la que disponemos de:

- Clases para la definición de los agentes involucrados:

```
<owl:Class rdf:ID="Citizen">
  <rdfs:subClassOf
rdf:resource="#Person"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty
rdf:resource="#hasNationality"/>
    </owl:Restriction>
  </rdfs:subClassOf>
```

- Eventos propios del ciclo de vida de los actores

```
<owl:Class rdf:ID="Birth">
  <rdfs:subClassOf
rdf:resource="#LifeEvent"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty
rdf:resource="#hasProtagonist"/>
<owl:toClass
 rdf:resource="#Person"/>
    </owl:Restriction>
  </rdfs:subClassOf>
```

  Esto nos permite además disponer un modo sencillo de una descripción de servicios mediante OWL-S de un modo sencillo y con una alta coherencia en el modelo global del sistema.

- Propiedades relevantes del dominio:

```
<owl:ObjectProperty
rdf:ID="chiperServize">
  <rdfs:subPropertyOf
rdf:resource="#hasSecurity"/>
  <rdfs:range
    rdf:resource="#Service"/>
</owl:ObjectProperty>
```

7.2. Comunicación

Los agentes involucrados en este intercambio de datos para la provisión de servicios deberán hacer uso de algún modelo que permita expresar y pedir la información necesaria en cada caso. Para

ello, se ha recurrido al diseño de una especificación XML mediante Schemas. El uso de este soporte nos permite hacer uso de tecnologías relacionadas para satisfacer necesidades del sistema: encriptación (SAML[12]), envío de información binaria, uso de redes de datos actual basadas en HTTP, ...

Estos mensajes expresados en XML estarán altamente relacionados con la ontología facilitada para describir el sistema. Hacer estos independientes resulta muy complejo ya que implicaría un nivel de inteligencia muy sofisticado difícil de alcanzar con el estado del arte actual en el campo de los agentes inteligentes.

7.3. Implementación de agentes

Dado el ciclo de vida que experimenta este tipo de sistemas (alta reutilización de funcionalidades desde diferentes puntos de una AP y alta velocidad de actualización), nos hemos decantado por el uso de un paradigma de programción COP(Component-Oriented Programming)[1] para el desarrollo de aplicaciones en lo que a los servidores se refiere. Los demás agentes del sistema, si bien tendrán una naturaleza heterogénea también está previsto su desarrollo mediante el uso de componentes reutilizables en el lado del cliente. Los motivos de esta elección están basados en la alta adecuación de este paradigma en lo que a reutilización y actualización de componentes de software se refiere. El desarrollo dentro de este modelo puede también apoyarse en los modernos IDEs que soportan de un modo adecuado el uso de COP en desarrollos de gran escala.

8. Conclusiones

En el momento actual se está produciendo una proliferación de soluciones dentro del campo del eGovernment. Desafortunadamente, estas soluciones abordan el problema desde un nivel de modelo de datos o funcional y no prevén problemas de interoperabilidad. En este artículo se hace una propuesta de una arquitectura de servicio en la que se identifican los agentes involucrados y soporta la prestación de servicios desde una perspectiva semántica. De este modo estamos en condiciones de proveer servicios desde un nivel superior y, así, estar en mejor disposición para la prestación de servicios de alto valor añadido, esencialmente mediante las funcionalidades implementadas en el *broker*, de un modo altamente interoperable, gracias al uso del nivel semántico en el sistema, y con facilidades específicas del dominio; referidas éstas últimas a temas de seguridad, control de información y notificaciones de servicios.

Un aspecto que se debe tener en cuenta es que independientemente de que la plataforma de servicio propuesta esté pensada para el uso de una base semántica permite que agentes que operen a un nivel más bajo puedan seguir prestando servicios en las mismas condiciones que lo hacen actualmente, es decir, el sistema es compatible con los sistemas actuales y también es posible el desarrollo de nuevos agentes en el sistema con capacidades limitadas para operar (sin soporte para descubrir e invocar servicios semánticamente) que, si bien no usando todo el potencial del sistema, sí podrán participar de un modo operativo.

El estado final del sistema se alcanzará cuando consigamos proporcionar una infraestructura en la que sea posible que cualquier entidad/grupo o sujeto sea capaz de desarrollar su propio agente y que este sea capaz de integrarse en la red semántica presentada y, por si mismo, realizar las operaciones establecidas en el sistema.

Agradecimientos

Queremos expresar nuestro agredecimiento al "Ministerio de Educación y Ciencia" por su apoyo parcial a este trabajo a través del proyecto "MetaLearn: metodologías, arquitecturas y lenguajes para servicios adaptativos para E-learning" (TIN2004-08367-C02-01).

Referencias

[1] Kai Qian Andy Ju An Wang. *Component-Oriented Programming.* Wiley, 2005.

[2] Len Bass, Paul Clements, and RickKazman. *Software Architecture in Practice.* Addison Wesley, 2003.

[3] Tim Berners-Lee, James Hendler, and Ora Lassila. The semantic web- a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. Web available, 2005. `http://www.scientificamerican.com/article.cfm?articleID=00048144-10D2-%1C70-84A9809EC588EF21\&catID=2`.

[4] Booch, Rumbaugh, and Jacobson. *The Unified Modeling Language User Guide.* Addison Wesley Professional, 199.

[5] OWL-S Coalition. Owl-s: Semantic markup for web services. Web available, 2005. `http://www.daml.org/services/owl-s/1.1/`.

[6] World Wide Web Consortium. Rdf. Web available, 2005. `http://www.w3c.org/RDF/`.

[7] French Government. Adea. Web available, 2004. `http://www.adae.gouv.fr/adele/`.

[8] T. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, pages 199–220, 1993.

[9] HP. Jena. 2005. http://www.hpl.hp.com/semweb/.

[10] KBSt. Saga. Web available, 2005. http://www.kbst.bund.de/-,182/SAGA.htm.

[11] United Nations. Benchmarking e-government: A global perspective. Web available, 2005. http://www.unpan.org/egovernment4.asp.

[12] Cover Pages OASIS. Security assertion markup language (saml). Web available, 2005. http://xml.coverpages.org/saml.html.

[13] Ontoknowledge project. Oil. Web available, 2005. http://www.ontoknowledge.org/oil/.

[14] SDK WSMO working group. Wsmo. Web available, 2005. http://www.wsmo.org/TR/d2/v1.1/.

[15] POPKIN Software. Feaf. Web available, 2004. http://government.popkin.com/frameworks/feaf.htm.

[16] Stanford Medical Informatics. Protege. Web available, 2005. http://protege.stanford.edu/.

[17] Stanford Medical Informatics. Protege owl plugin. Web available, 2005. http://protege.stanford.edu/plugins/owl/.

[18] The DARPA Agent Markup Language Homepage. Daml. Web available, 2005. http://www.daml.org/.

[19] UK GovTalk. e-GIF. Web available, 2004. http://www.govtalk.gov.uk/.

[20] W3C. Web ontology language. Web available, 2004. http://www.w3.org/2004/OWL/.

[21] W3C. Web services architecture. Web available, 2005. http://www.w3.org/TR/ws-arch/.

Jornadas de seguimiento

Proyectos del Plan Nacional de I+D+I
Programa de Tecnología de Servicios para la Sociedad de la Información

Editores

José J. Pazos Arias Javier Aracil Rico

UNIVERSIDAD DE VIGO

MINISTERIO DE
EDUCACIÓN Y CIENCIA

AgentWeb: Agentes y Ontologías para la Gestión de Derechos Digitales y Servicios Web

Jaime Delgado
Grupo de Aplicaciones Multimedia Distribuidas (DMAG)
Departamento de Tecnología, Universitat Pompeu Fabra
Passeig de la Circumval·lació, 8, 08003 – Barcelona
Teléfono: 93 542 22 55 Fax: 93 542 25 17
E-mail: jaime.delgado@upf.edu

Abstract. *El proyecto AgentWeb (Agentes y Ontologías para la Gestión de Derechos Digitales y Servicios Web), TIC2002-01336, empezó en Diciembre de 2002 y está previsto termine en Noviembre de 2005. Está siendo realizado por el grupo de Aplicaciones Multimedia Distribuidas (DMAG, Distributed Multimedia Applications Group), del Departamento de Tecnología de la Universitat Pompeu Fabra en Barcelona. Se cubren varias áreas temáticas, dividiéndose el trabajo en una serie de tareas. Este breve resumen hace especial hincapié en aquellas actividades más relacionadas con la Gestión de Derechos Digitales (DRM, Digital Rights Management), que son las que están teniendo más impacto hacia el exterior. En este contexto, cabe destacar que muchos de los resultados obtenidos en este proyecto han sido contribuidos a la estandarización internacional donde han sido verificados y aceptados.*

1 Gestión de Derechos Digitales

Entre las iniciativas internacionales centradas en la gestión y distribución de contenidos multimedia protegidos y gobernados, podemos destacar el estándar MPEG-21, que pretende especificar un marco de trabajo para la definición y uso de contenidos digitales. Dentro de las numerosas partes del MPEG-21 (cerca de 20 en este momento), las más relacionadas con nuestro trabajo en AgentWeb son aquéllas que tratan de los llamados items digitales (estructuras de datos que incluyen los contenidos multimedia y toda la información necesaria para su gestión), y las que tratan de la gestión y protección de derechos digitales. Estas últimas incluyen un lenguaje de expresión de derechos (REL, Rights Expression Language), un diccionario de datos (RDD, Rights Data Dictionary) y la posibilidad de expresar información de protección de los items digitales, lo que se llama IPMP (Intellectual Property Management and Protection).

Antes del inicio del proyecto, ya habíamos empezado a desarrollar una ontología de derechos de propiedad intelectual, que llamamos IPROnto [1], que se ha ido refinando y conectando con las iniciativas actuales de estandarización. En concreto, estamos especificando nuevas ontologías para MPEG-21 y para otro lenguaje de expresión de derechos, ODRL (Open Digital Rights Language).

Estas iniciativas tienen una aproximación sintáctica, al contrario que IPROnto, cuyo enfoque es semántico. Se basan fundamentalmente en definir la gramática de un REL. Este aspecto se implementa utilizando XML Schema. Por otra parte, es necesario dar un significado a los elementos definidos en la gramática para establecer una interpretación estándar del lenguaje. Esto se hace en los diccionarios de datos de derechos, que en MPEG-21 son una parte diferente al REL (el RDD) y en ODRL está integrado con el mismo lenguaje.

Con la intención de integrar la iniciativas presentadas (ODRL y MPEG-21) con IPROnto, ambas han sido traducidas al lenguaje común de creación de ontologías Web OWL. Este lenguaje es suficientemente potente como para poder representar las diferentes especificaciones de estas iniciativas, ya sean XML Schemas o el lenguaje utilizado en RDD. En el primer caso, se han mapeado las construcciones propias a construcciones de OWL. Por lo que respecta a MPEG-21 RDD, en este caso sí que se trata de una ontología. Lo que se ha hecho es mapear las construcciones RDD a las construcciones de OWL. Con esto se consigue una ontología del MPEG-21 RDD en OWL. Se pueden encontrar más detalles en [1].

Además de nuestro trabajo de formalización, hemos hecho un trabajo de verificación e implementación de algunas partes del MPEG-21. Desde mediados de 2003, hemos venido participando en y contribuido a las reuniones de MPEG, un complejo grupo de expertos (entre 300 y 400 personas activas) que está desarrollando los diferentes estándares MPEG, desde los antiguos MPEG-1 y 2, hasta el MPEG-4, el MPEG-7 y, sobre todo actualmente, el MPEG-21.

Nuestras primeras contribuciones, implementaciones del lenguaje de expresión de derechos (REL), fueron aceptadas y actualmente están incluidas en la Parte 8 del estándar (MPEG-21 Reference Software). Éstas incluyen herramientas básicas para analizar y crear licencias, algoritmos de autorización, acceso a

diccionarios de datos, etc. Después, hemos continuado trabajando en esa línea haciendo desarrollo de software de referencia para las distintas partes del estándar, lo que ha implicado trabajar en temas de investigación relacionados, como la validación de las expresiones de derechos, para lo que se han definido una serie de reglas que determinan si una licencia es válida según el estándar MPEG-21 REL, y la autorización de acceso a contenido gobernado y/o protegido. Hemos definido también el formato de los datos para los diferentes elementos que forman el modelo de autorización MPEG-21 REL.

Últimamente estamos trabajando en nuevas líneas, como en la parte 4 del estándar, donde realizamos en su momento una contribución a la "llamada de propuestas" de IPMP que consistía en la definición de un esquema XML para representar la información IPMP y su asociación con items digitales y expresiones de derechos. La evaluación realizada por el comité pertinente seleccionó unas pocas de las varias tecnologías propuestas, incluida la nuestra. Como resultado, uno de los miembros del DMAG es co-editora de este estándar internacional en desarrollo.

Finalmente, hemos empezado a trabajar en el problema de la interoperabilidad entre los sistemas de gestión de derechos digitales, y en particular por lo que respecta a los lenguajes de expresión de derechos. En concreto, hemos propuesto cómo adaptar el REL de MPEG-21 para ser compatible con algunos subconjuntos de ODRL que han propuesto importantes foros industriales, como la Open Mobile Alliance (OMA). Asimismo, se han modificado algunas de las herramientas software desarrolladas inicialmente para MPEG-21 para que también funcionen con ODRL. Mencionar finalmente que otras iniciativas internacionales, como el Digital Media Project (DMP) nos han solicitado utilizar nuestras contribuciones en el área de interoperabilidad.

2 Otras líneas de investigación

2.1 Arquitectura de agentes para DRM

Se ha desarrollado una arquitectura de agentes basada en JADE-LEAP que permite llevar a cabo la negociación automática de derechos. Para modelar el comportamiento de los agentes se ha utilizado el sistema experto Jess (Java Expert System Shell) como lenguaje para la expresión de las reglas que controlan los aspectos dinámicos de los agentes y su comportamiento. Las reglas se complementan con conocimiento del dominio DRM adquirido de la ontología IPROnto. Asimismo, se ha desarrollado una herramienta que permite extraer contenido semántico de grandes volúmenes de datos.

2.2 Interoperabilidad de metadatos

En las aplicaciones de búsqueda y metabúsqueda es habitual que se vean involucrados agentes que usan diferentes esquemas de metadatos para definir y calificar los recursos ofrecidos/buscados. Por tanto, la interoperabilidad tanto a nivel sintáctico como semántico entre diferentes esquemas es un aspecto clave en este tipo de aplicaciones. En esta línea, hemos definido un modelo de interoperabilidad de metadatos a nivel semántico cuya principal característica es que no impone el uso de ningún esquema de metadatos concreto a los diferentes actores del sistema.

2.3 Control de flujo de servicios

Se han definido tipos de servicios y sus flujos de trabajo asociados y se ha definido el ciclo de vida de un contenido multimedia al que se quiere dotar de protección mediante derechos digitales. Para ello, se ha utilizando la idea de la descripción de flujos de trabajo ya introducida para los servicios.

2.4 Servicios de seguridad

Una de las líneas de trabajo de esta tarea se centra en la protección de información multimedia representada en XML. En concreto, se han investigado mecanismos avanzados para asegurar la autenticidad e integridad de las licencias de distribución de contenidos multimedia expresadas con lenguajes de derechos. Todos los lenguajes de este tipo actualmente en uso se basan en XML, por lo que la técnica más apropiada consiste en aplicar la norma XML Signature para firmar digitalmente las licencias. En este caso, el primer paso para la generación de la firma siempre es la canonización de la información XML, para lo que el W3C (World Wide Web Consortium) publicó la especificación "Canonical XML". Sin embargo, se puede obtener mayor flexibilidad si se aplica una canonización a nivel semántico, en lugar de únicamente "Canonical XML" que define transformaciones a nivel de sintaxis XML.

3 Conclusiones

Los resultados de los trabajos de investigación indicados en este breve resumen han dado lugar, además de a las contribuciones y estándares apuntados, a tres Tesis Doctorales (y una más a punto de presentarse) y a cerca de 50 publicaciones internacionales. Todo ello se puede encontrar en la Web del grupo [1].

Referencias

[1] Distributed Multimedia Applications Group (DMAG): http://dmag.upf.edu

Proyecto ARCADIA: Generación automática y rediseño de documentos web en un sistema de adquisición de conocimientos colaborativo, autónomo, distribuido e interactivo

Xavier Alamán, Manuel Alfonseca, Pablo Castells, Ruth Cobos, Fernando Díez, Juan de Lara, José Antonio Macías, Jaime Moreno, Roberto Moriyón, Alfonso Ortega, Álvaro Ortigosa, Eduardo Pérez, Estrella Pulido, Francisco Saiz, Leila Shafti, Juan Alberto Sigüenza
Departamento de Ingeniería Informática, Universidad Autónoma de Madrid
Escuela Politécnica Superior. C/ Tomás y Valiente, 11. Campus de Cantoblanco. .
28049 Madrid
Teléfono: 91 497 2278. Fax: 91 497 2235
E-mail: Manuel.Alfonseca@uam.es

Abstract. *The objective of project ARCADIA (Automatic generation and Redesign of web documents in a Collaborative, Autonomous, Distributed, and Interactive knowledge Acquisition system) is the development of integrated techniques and tools for user communities in the web to provide a semi-authomatized organization of knowledge as a consequence of collaborative participation mechanisms, with a high degree of freedom in the representation of knowledge, and providing personalized means to consult that knowledge by means of user-adapted documents.*

El objetivo del proyecto ARCADIA (Automatic generation and Redesign of web documents in a Collaborative, Autonomous, Distributed, and Interactive knowledge Acquisition system) es el desarrollo de técnicas y herramientas integradas de gestión del conocimiento para comunidades de usuarios en la web que permitan:

a) La organización semi-automatizada del conocimiento en base a mecanismos de participación colaborativa.
b) Un alto grado de libertad para definir la estructura a utilizar en la representación del conocimiento.
c) La consulta personalizada del conocimiento mediante la generación automática de documentos adaptados al usuario y otras condiciones de uso.

Los distintos módulos del entorno funcionan sobre una base de conocimientos definida mediante una ontología, en la que se capturan los conceptos y categorías semánticas mediante los que se describe el conocimiento de un dominio dado. La ontología sirve de vehículo para la compartición y reutilización de conocimiento entre los módulos del proyecto. El trabajo realizado se basa en la perspectiva de la denominada web semántica, que propone una web mejor organizada y estructurada que la actual, mediante la introducción de conocimiento semántico explícito sobre los recursos disponibles en la red, para facilitar la localización, compartición e integración de los mismos. Las técnicas desarrolladas en el proyecto se han validado mediante la implementación de dos aplicaciones.

El proyecto se ha plasmado en la realización de las siguientes tareas:

1. Análisis de requisitos: revisión del estado del arte en tecnologías de la web semántica, con atención a lenguajes y estándares de definición de ontologías, entornos de desarrollo y gestión y metodologías de construcción, integración, transformación y evolución de ontologías, visualización, navegación y generación de portales web, servicios disponibles en la web semántica y aplicaciones existentes.
2. Diseño de un lenguaje para la representación del conocimiento: se seleccionó una metodología para la creación de ontologías diseñada por la universidad de Stanford y un módulo que permite la creación automática de casos de ontologías a partir de una descripción de la correspondencia entre las estructuras de las bases de datos de partida (bases documentales y relacionales, cuyo volumen hace inviable la creación manual de meta-datos o instancias de la ontología que describan la información) y los elementos de la ontología de destino. Además, partiendo de una ontología de conocimientos de

matemáticas, se ha diseñado una herramienta para construir materiales interactivos en base a los mismos y se ha realizado un trabajo similar para la creación de ejercicios interactivos a partir de colecciones de textos con información semántica interna que forman parte de cursos. Se ha desarrollado un sistema basado en anotaciones semánticas asociadas a partes de documentos, que permite añadir interactividad a documentos estáticos. El sistema está basado en un formalismo genérico de asociación de anotaciones semánticas a partes de documentos, la especificación de un conjunto de semánticas específicas que representan distintos tipos de interactividad sobre las partes asociadas, herramientas de autor para el diseño interactivo de documentos interactivos de distintos tipos, con sistemas avanzados de ejercicios para el reforzamiento del aprendizaje conceptual, y entornos de ejecución interactiva de los documentos anteriores. Este trabajo ha dado lugar a una patente mundial en tramitación. También se ha trabajado en el diseño de otras formas de utilización de información semántica contenida en distintos documentos para la definición de aplicaciones colaborativas, en el ámbito del aprendizaje colaborativo CSCL.

3. Diseño e implementación de un módulo de adquisición del conocimiento: se empezó por un conjunto de casos de uso más relevantes para una implementación restringida al ámbito académico. Como resultado, se obtuvieron subsistemas pare el diseño y servicio, las clases del diseño, las realizaciones de casos de uso-diseño y la vista arquitectónica del modelo de diseño. También se ha profundizado en la forma en que los requisitos ampliados, junto con las limitaciones del entorno de implementación, influyen en el módulo dentro del sistema completo. Se ha obtenido una versión refinada de los subsistemas de diseño y de servicio, las clases de diseño, las realizaciones de casos de uso-diseño y la vista arquitectónica del modelo de diseño. Asímismo, se ha realizado el diseño de la ontología que soportará el conocimiento que maneja el módulo y permite su adquisición, enriquecimiento y cristalización. En cuanto a la implementación, se han implementado todos los casos de uso considerados en un ámbito de aplicación general. El módulo implementado incorpora un núcleo para la gestión del conocimiento, un sub-módulo de monitorización y análisis de la actividad de los usuarios, y un sub-módulo de análisis y enriquecimiento.
4. Modelado e implementación de un sistema de presentación del conocimiento: se ha diseñado un lenguaje y una arquitectura para la visualización y navegación en bases de conocimiento basadas en ontologías mediante documentos virtuales. Se ha implementado un prototipo del módulo de visualización y una herramienta de autor inteligente asociada, para la edición interactiva de los documentos virtuales (páginas web dinámicas). También se ha completado un módulo de visualización y navegación semántica para bases de conocimiento basadas en ontologías.
5. Diseño e implementación del módulo de análisis automático de datos y modelado de usuario: se ha diseñado un método que combina varias técnicas de minería de datos y aprendizaje automático, como la inducción constructiva y los algoritmos genéticos. Este método analiza los datos para encontrar relaciones complejas entre ellos, y genera una nueva representación de datos basada en nuevos conceptos intermedios, propuestos por la propia herramienta, para facilitar la extracción final del conocimiento. Además, esta nueva representación muestra las relaciones entre los datos de forma transparente, más fácil de comprender. Se ha implementado un primer prototipo de dicho método, que se ha evaluado, con resultados provisionales satisfactorios, utilizando datos sintéticos.
6. Integración y validación: se está completando la integración del módulo de visualización con las ontologías y bases de conocimiento desarrolladas por otros módulos.
7. Desarrollo de una aplicación: se han realizado dos casos de estudio, con sus ontologías y bases de conocimiento correspondientes, que se alimentan automáticamente a partir de bases documentales y de datos proporcionadas por las empresas colaboradoras: una plataforma para la gestión de la hemeroteca del diario SEGRE S.L.U., y otra para la gestión de información económica y financiera en colaboración con Tecnología, Información y Finanzas (Grupo Analistas). Estos dos casos de estudio han dado lugar a la realización de sendos prototipos, públicamente disponibles en la red.

Aproximación Sistemática al Desarrollo e Integración de Archivos Digitales en Web

Proyecto[1] TIC2002-04050-C02-01
Universidad Rey Juan Carlos
Investigadora Principal: Esperanza Marcos
esperanza.marcos@urjc.es

Investigadores: C.J. Acuña, M. V. de Castro, P. Cáceres, J. M. Cavero, B. Vela
Grupo de Investigación KYBELE

1 Introducción

El proyecto que se presenta tenía como objetivo principal la especificación de un marco metodológico que facilitara el desarrollo sistemático de un tipo concreto de Sistemas de Información (SI): portales Web de acceso integrado a fuentes de datos con contenido digital [11]. Para ello, se han llevado a cabo las siguientes actividades: a) especificación de un marco metodológico para el desarrollo de Sistemas de Información Web (SIW); b) especificación de una arquitectura de integración de portales Web; c) aplicación del marco y de la arquitectura a casos reales.

2 MIDAS: Marco para el Desarrollo Sistemático de SIW

El marco metodológico propuesto, denominado MIDAS [19], define una arquitectura dirigida por modelos [3] y un proceso de desarrollo ágil [4][5]. La arquitectura, considera dos dimensiones ortogonales. La primera se basa en la arquitectura MDA: Modelos Independientes de Computación (CIMs), Modelos Independientes de Plataforma (PIMs) y Modelos eSpecíficos de Plataforma (PSMs). La segunda dimensión considera los diferentes aspectos que han de ser tenidos en cuenta en el desarrollo de un SIW; en MIDAS, y como punto de partida, se han tenido en cuenta los aspectos de hipertexto, contenido y comportamiento. Esta arquitectura cuenta con la ventaja de ser fácilmente escalable, permitiendo la incorporación, siempre que así se considere, de un nuevo aspecto o de nuevos modelos.

En general, se propone la utilización de UML, por lo que se han definido los perfiles UML necesarios para la especificación, tanto de los modelos, como de las transformaciones entre ellos. En el aspecto de contenido se ha definido un perfil para Bases de Datos (BD) Objeto-Relacionales [8][9][10] y otro para BD XML [16][17][18]. Respecto al aspecto del hipertexto, se ha especificado un método orientado a servicios de usuario que facilita la obtención de una interfaz de navegación Web más intuitiva y fácilmente navegable para el usuario [6][7][8][14]. En el aspecto de comportamiento, se ha comenzado por la especificación de un perfil para el modelado de servicios Web basado en WSDL [7]. Así mismo, se está trabajando en la especificación de modelos y transformaciones, para el modelado de la composición de servicios Web (basados en BPEL).

En la actualidad, y en colaboración con el prof. R. Wieringa de la Universidad de Twente (Holanda), donde M. V. de Castro ha realizado una estancia, se están definiendo los modelos del nivel CIM así como las transformaciones entre éstos y los modelos del nivel PIM. En paralelo, se ha comenzado con el desarrollo de una herramienta CASE que soporte el desarrollo basado en MIDAS de SIW. Se ha especificado la arquitectura de la herramienta, el repositorio de la misma, así como el módulo para la generación automática de WSDL.

3 Arquitectura de Integración de Portales

Aunque la propuesta inicial del proyecto planteaba una arquitectura de integración de fuentes de datos, ya en los inicios del mismo nos dimos cuenta de las ventajas que podía suponer abordar la integración de portales, al permitir integrar datos sin necesidad de acceder al repositorio de almacenamiento de los mismos. Por otra parte, la integración de portales nos permite integrar, además de fuentes de datos, servicios. Por todo ello, comenzamos con la especificación de una arquitectura de integración de fuentes (tanto datos como servicios) de portales Web. Dicha arquitectura, basada en servicios Web, se comenzó a definir en colaboración con el grupo STORM del CNRS (Francia) durante la estancia de J.M.Cavero.

En la actualidad, se está implementando dicha arquitectura utilizando la plataforma WISMO. Este trabajo se está llevando a cabo en colaboración con el grupo del prof. C. Bussler del DERI (Irlanda), donde C.J.Acuña ha realizado una estancia.

---

[1] Proyecto Coordinado TIC2002-04050-C02. Participantes: URJC, UPM. Coordinación: Esperanza Marcos (URJC)

4 Aplicaciones

Se han definido distintos SIW, que integran fuentes con información digital, para validar y refinar tanto la arquitectura de integración como las propuestas metodológicas. Entre ellos, cabe destacar la implementación de un portal Web para la gestión y procesamiento de imágenes médicas [1][2]. En la actualidad se está trabajando en la implementación de un portal de integración de este tipo de información. El desarrollo se ha llevado a cabo utilizando Oracle y .NET y en colaboración con el grupo de Grupo de Tecnología Electrónica, Bioingeniería e Imagen Médica (GTEBIM) de la URJC, dirigido por el Dr. J. A. Hernández.

5 Conclusiones

En resumen, podemos decir que se ha obtenido un marco metodológico (que integra técnicas de desarrollo ágil en una arquitectura basada en modelos) para el desarrollo de SIW y una arquitectura para la integración de fuentes. Ambos se están validando mediante su aplicación a distintos SIW reales. Se cuenta con un prototipo en funcionamiento del portal para la gestión y procesamiento de imágenes médicas. Además, se tiene un primer prototipo de una herramienta CASE que permita el desarrollo basado en el marco metodológico propuesto.

El proyecto ha dado (y está dando) lugar a numerosas publicaciones tanto nacionales como intencionales en foros de reconocido prestigio. Además, en el marco del mismo se ha leído una Tesis doctoral, hay otra en fase de finalización (se estima su defensa en noviembre de 2005) y otras 6 en proceso de realización.

Referencias

[1] C. Acuña, E. Marcos, V. de Castro y J.A. Hernández. Managing Medical Images. *ERCIM News*. Ed. ERCIM EEIG, Francia. No. 58, pp. 54-55. Julio, 2004.

[2] C. Acuña, E. Marcos, V. de Castro y J.A. Hernández. A WIS for Medical Images Management. *International Symposium on Biological and Medical Data Analysis (ISBMDA'04)*. Ed. J. M. Barreiro, F. Martín-Sánchez, V. Maojo y F. Sanz. Springer Verlag, LNCS, 2004, pp.49-59.

[3] P. Cáceres, E. Marcos y B. Vela. "A MDA-Based Approach for Web Information System Development". *Workshop in Software Model Engineering (WiSME'03)*. UML Forum. USA, 2003.

[4] P. Cáceres, E. Marcos y F. Díaz. Agile Model Driven Development in WIS: A Case Study. *Advances in Theory, Practice and Education (ISD'04)*. Eds. Vasilecas, Caplinskas, Wojtkowski, Wojtkowski, Zupancic, Wrycza. 2004, pp. 341-351.

[5] P. Cáceres, F. Díaz y E. Marcos. Integrating an Agile Process in a Model Driven Architecture. Actas del First *Workshop on Web Applications and Middleware (ULM'04)*. Ed. Dadam, Reichert, 2004, pp. 265-270.

[6] P. Cáceres, E. Marcos y V. de Castro. Navigation Modeling from a User Services Oriented Approach. *Advanced en Information Systems (ADVIS'04)*. Ed. Tatyana Yakhno, 2004, pp. 150-160.

[7] P. Caceres, E. Marcos y V. de Castro. Modelado Navegacional desde una Perspectiva Orientada a Servicios de Usuario. *IEEE América Latina*. Aceptado para publicación (seleccionado de las JISBD'04).

[8] V. de Castro, E. Marcos, P. Cáceres. A User Services Oriented Method to model WIS. *5th International Conference on Web Information Systems Engineerin (WISE'04)*, Ed. Zhou, Su, Papazoglou, Orlowska, Jeffery. Springer Verlag, LNCS, pp.41-52, 2004.

[9] J.M. Cavero, E. Marcos, B.Vela y C. Costilla. "Modelling ORDB Queries using UML". Actas de la *Internacional Conference on Enterprise Information Systems (ICEIS'03)*. Ed. O. Camp, J. Filipe, S. Hammoudi, M. Piattini. Vol. 3, pp.535-539, 2003.

[10] J.M. Cavero, E. Marcos, B. Vela, P. Cáceres y C. Costilla. Integrating Query and Hypertext Modeling in WIS Development., IADIS International Conference. Vol. II, Ed. Isaias y Karmakar, 2004, pp-1195-1198.

[11] E. Marcos, P. Cáceres, J. M. Cavero, B.Vela, C. Costilla, S. Eibe, E. Menansalvas y J. Sáenz. "DAWIS: Sistematización del Desarrollo de Portales para el Acceso Integrado a Archivos Digitales en la Web". *Taller de Integración semántica de fuentes de datos distribuidas y heterogéneas (VII JISBD)*. El Escorial (Madrid), 19-21 noviembre 2002.

[12] E. Marcos, V. de Castro, B. Vela. Representing Web Services with UML: A Case Study. En actas de *la International Conference on Service-Oriented Computing – ICSOC 2003*. Ed.: M. E. Orlowska, S. Weerawarana, M. P. Papazoglou y J. Yang. Springer Verlag, LNCS-2910, pp.15-27, 2003.

[13] E. Marcos, B. Vela, J.M. Cavero. A Methodological Approach for Object-Relational Database Design using UML. *Informatik Forschung und Entwicklung*. Ed. Springer Verlag, Heidelberg (Alemania), Vol: 18: 3-4, pp. 152-164, Abril, 2004.

[14] E. Marcos, P. Cáceres, V. de Castro. An approach for Navigation Model Construction from the Use Case Model. En *Knowledge and Model Driven Information Systems Engineering for Networked Organization (CAISE'04 Forum Proceedings)*. Ed. J. Grabis, A. Persson y J. Stirna, 2004, pp.83-92.

[15] B. Vela y E. Marcos. "Una extensión de UML para representar XML Schemas". Memorias del *6º Workshop Iberoamericano de Ingeniería de Requisitos y Ambientes Software* (Ideas 2003), Ed. M. Piattini, L. Cernuzzi y F. Ruíz. 2003, pp.109-119.

[16] B. Vela y E. Marcos. "Extending UML to represent XML Schemas". *The 15th Conference on Advanced Information Systems Engineering. CAISE'03 Forum*. Proceedings Short Papers. Ed. J. Eder y T. Welter. 2003, pp.97-100, 2003.

[17] B. Vela Sánchez, C. J. Acuña, y E. Marcos. Una aproximación dirigida por modelos para el desarrollo de bases de datos XML. *IX JISBD*. Ed. J. Hernández, E. Pimentel. pp.145-158, 2004.

[18] B. Vela, C. Acuña, E. Marcos. A Model Driven Approach for XML Database Development. *International Conference on Conceptual Modeling (ER'04)*. Ed. P. Atzeni, W. Chu, H. Lu, S. Zhou y T. Wang. Springer Verlag, LNCS, 2004, pp.780-794.

[19] B. Vela, P. Cáceres, V. Castro y E. Marcos. MIDAS: Una Aproximación Dirigida por Modelos para el Desarrollo Ágil de SIW. Aceptado para su publicación en: *Ingeniería Web y Patrones de Diseño*. Ed. P. Díaz, S. Montero e I. Aedo. Prentice-Hall.

The National Research Project: '*Digital Archive Web Information Systems*'. Summary and Lessons Learned

Carmen Costilla Rodríguez
Inv. Ppal. del Proyecto TIC02-04050-02-2 (2002-2005, TIC National Programme)
Dep. Ingeniería de Sistemas Telemáticos, Universidad Politécnica de Madrid
ETSI de Telecomunicación, Ciudad Universitaria s/n, 28040 – Madrid (Spain)
http://sinbad.dit.upm.es, costilla@dit.upm.es

*A **virtual** (non-materialized) **integrated Web reference architecture** to query multiple Digital Archives (**DA**) has been defined. For each concise integration, this architecture is dynamically generated from the web user query and the many participant DA, without affecting each local Web site organization that remains unchanged at each original DA site. Ontologies and XML are the main used technology.*

*The following goals have been overcome: a) The **UML** specification of a **Digital Archive Conceptual Model** (applied to local DA), providing four Archival **Ontologies OWL** coded; b) A **Data Extractor Model** definition, made by 'wrappers' hiding the data sources heterogeneity; c) The mediator specification, providing a **Unified Archival Semantic Web**, merging ontologies through mappings and repositories; d) A **Web Service Architecture** proposal, web-centric and J2EE compliant.*

*We have started from the DA management system, successfully running at the '**Parliament of the Asamblea de Madrid**', we have built (1997-2000). Additionally, the **CRC Information Technologies** Spanish enterprise help us to verify these project results. During these first 29 months of this project, a total of 25 papers dealing with these research topics have been published on several International Congresses and we had been involved in different Scientific Committees.*

Keywords: Web Digital Archives, Web Information Systems, e-government, Virtual and Dynamic Integrated Web Architecture, Mediators and Wrappers, Semantic Web, Ontologies, OWL, Heterogeneous Web Data Source Extraction, XML.

1 Introducción

Este proyecto investiga el acceso web integrado a múltiples archivos digitales (en adelante, AD). El AD es una ingente colección de documentos multimedia e información descriptiva (metadatos) almacenada en repositorios (datos documentales, estructurados y multimedia). La digitalización y la integración de múltiples AD debe permitir que el usuario web acceda -con una simple consulta- a una información global provinente de diversos AD heterogéneos, virtualmente integrados y físicamente distribuidos en cualquier lugar del mundo.

Hemos partido de la experiencia adquirida en dos líneas de trabajo recientes: 1) el Sistema de Gestión Parlamentario ***SIAP***[1], que opera con éxito en la Asamblea de Madrid desde 1999. Su subsistema de Gestión del Archivo Parlamentario constituye aquí una valiosa contribución de partida. 2) ***EDAD-UPM***[2], proyecto complementario al que ahora nos ocupa.

Investigamos la arquitectura integradora web (con 'mediator' y 'wrappers') entre las fuentes de datos heterogéneas y el 'browser' del usuario para permitir el acceso virtual, flexible y dinámico a múltiples AD. La arquitectura es escalable y a la medida del perfil del usuario web. La semántica de integración se basa en ontologías y es conforme a los estándares archivísticos (ISAD, ISAAR, EAD, DCMI, etc).

[1] SIAP *Sistema de Información para Ayuntamientos y Parlamentos* de 'CRC Information Technologies'

[2] EDAD-UPM *Entorno de Desarrollo de Archivos Digitales*. Proyecto financiado por la Comunidad de Madrid (2003-4)

2 Resultados Alcanzados

Nuestro trabajo se ha visto refrendado en foros internacionales, cuyos resultados científicos más relevantes se resumen a continuación.

2.1 Sobre la Arquitectura Web Integradora

Hemos especificado una **arquitectura web de referencia** para el acceso consultivo e integrado a diversos AD, como muestra la fig. 1. Dimos pasos importantes en aspectos fundamentales, como son: 1) especificación del modelo del AD, 2) extractores de datos y 'wrappers', que ocultan la heterogeneidad de las fuentes locales, 3) 'mediator' con ontologías, repositorios y mappings de la Web Semántica. La arquitectura permite cualquier integración de AD web. La consulta web hace que el 'Mediator' **genere el esquema global, las ontologías y mappings** 'al_vuelo', considerando los esquemas XML locales y la integración específica donde participa cada AD. Esto supone definir ontologías específicas para los AD y facilitar el intercambio de información entre cada ontología y los 'wrappers' de cada integración.

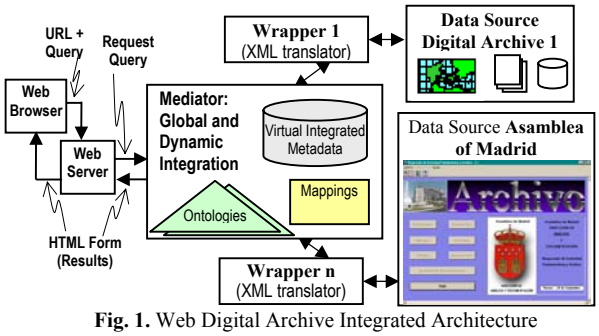


Fig. 1. Web Digital Archive Integrated Architecture

Así, los 'wrappers' también se generan dinámicamente a partir de la consulta y de los AD concretos participantes en cada integración.

2.2 Unificación Semántica Ontológica.

Hemos especificado una ***Web Semántica Global y Unificada*** para los AD, como muestra la fig. 2.

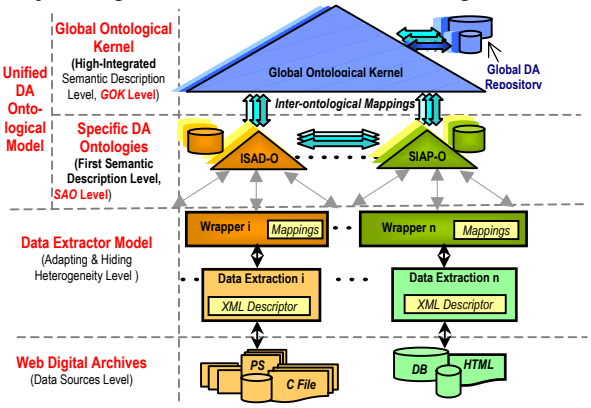


Fig. 2. The Proposed Virtual Integration Model

2.3 Modelo Extractor de Datos Heterogéneos

Hemos definido un ***Modelo Extractor de Datos Web Heterogéneos*** para cada posible AD participante, como muestra la fig. 3.

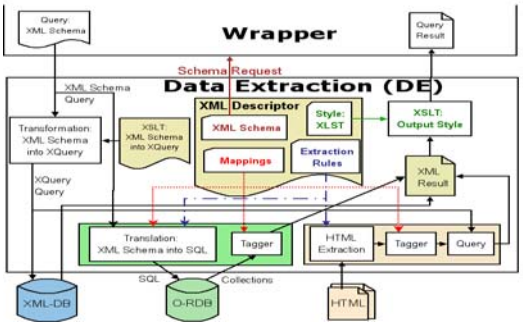


Figure 3. *DE* operating architecture.

2.4 Sobre la Definición de Servicios Web

La resultante es una colección de Servicios Web (bibliotecas Java) web-centric (J2EE compliant), junto a la proposición de técnicas y guías para automatizar, en lo posible, todo este proceso. Hemos empezado a definir **Servicios Web y Servicios Grid Semánticos**, como muestra la fig.4.

2.5 Resultados Científicos Publicados

1 Costilla C, Rodríguez M, Palacios J, Cremades J, Calleja A, Fernández R, *A Contribution to Web Digital Archive Integration from the Parliamentary Management System 'SIAP'*, **Frontiers in Artificial Intelligence and Applications**, **Data Bases and Information Systems. Selected papers from the Sixth Int. Baltic Conf.** (**DB&IS'2004**), Vol. 118, Barzdins J. and Caplinskas A. (eds), ISBN:1-58603-485-5, **IOS Press**, The Netherlands, pp. 273-287, 2005

2 de Miguel J, Calleja A, Costilla C, García M, *Web Services for a Semantic Web Integrated Architecture*, sent to 2005 **IEEE Int. Conf. Services Computing** (SCC'05), Orlando, USA, July 2005

3 Palacios J, Cremades J, Costilla C., *Towards a Web Digital Archive Ontological Unification,* Int. Conf. Information Technology and Applications (ICITA'05), **Agent, Data Mining and Ontologies, ADO'05**, **IEEE Computer Society**, Los Alamitos, California, USA, (accepted), Sydney, Australia, pp.x-yy, July 2005

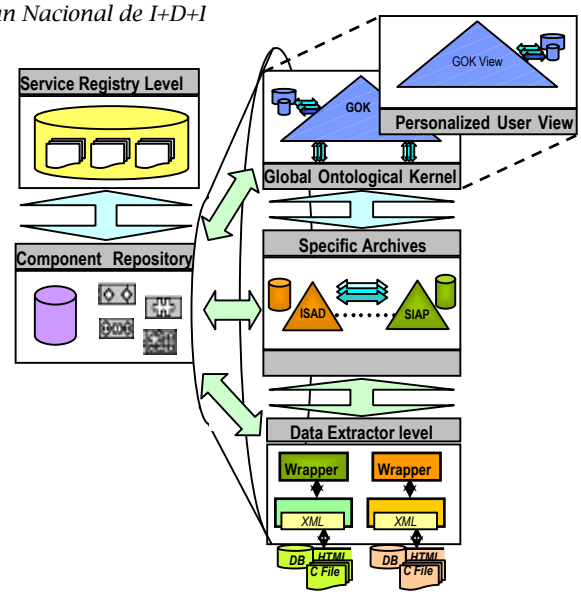


Fig. 4. Web Service Levels

4 Calleja A, Rodríguez M, Costilla C, Fernández R., *A Grid Semantic Approach for a Digital Archive Integrated Architecture,* Int. Conf. Information Technology and Applications (ICITA'05), **Agent, Data Mining and Ontologies, ADO'05**, **IEEE Computer Society**, Los Alamitos, California, USA, (accepted), Sydney, Australia, pp. xx-yy , July 2005

5 Costilla C, Cremades J, Fernández, R Calleja A, Palacios J, *e-government: Iniciativas Legislativas en la Web Semántica*, **X Congreso Nac. de Internet, Telecomunicaciones y Movilidad** (accepted), **Internet 2005,** Dep. Legal: xx, Palacio Congresos, **Asociación Usuarios de Internet,** Madrid, pp.xx-yy, April 2005

6 Costilla C, Palacios J, Cremades J, Vila J, *e-government: A Legislative Ontology for the 'SIAP' Parliamentary Management System*, E-Government: Towards Electronic Democracy, Proc. Int. Conf. **TCGOV 05**, **Lecture Notes in Artificial Intelligence**, LNAI 3416-0134, LNCS Series, ISBN 3-540-25016-6, **Springer** Berlin, Germany, pp. 134-146, IFIP 2005, March 2005

7 Thiran Ph, Jan van Heuvel W, Costilla C, Henrard J, Kabish T and Risch T., *Report on the Workshop on Wrapper Techniques for Legacy Databases*, **ACM Sigmod Record**, 34(2), (accepted), New York, USA, pp. xx-yy, June 2005

8 Illarramendi A, Marcos E and Costilla C., *RedBD: the Database Research Community in Spain,* **ACM Sigmod Record**, 34(1), New York, USA, pp.51-56, March 2005

9 Vila J, Costilla C., *Heterogeneous Data Extraction in XML,* First Int. Workshop on **Wrapper Techniques for Legacy Systems** (WRAP 2004), 11th W. Conf. Reverse Engineering (WCRE 2004), Proc. of **WRAP 2004,** ISSN: 0926-4515, **Tech. Univ. Eindhoven,** Delft, Eindhoven, pp. 1-15, Nov. 2004

10 Cavero J, Marcos E, Vela B, Cáceres P, Costilla C, *Integrating Query and Hypertext Modeling in Web Information Systems Development*. IADIS Int. Conf. **WWW/Internet 2004**, Madrid, Spain, 6-9 Octubre 2004.

11 Costilla C, Palacios J, Rodríguez M, Cremades J, Calleja A, Fernández R, Vila J, *Semantic Web Digital Archive Integration***,** 15th Int. Workshop on Database and Expert Systems Applications, Int. W. on **Web Semantics** (WebS 2004), Proc. **DEXA 2004**, ISBN: 0-7695-2195-9, **IEEE Computer Society** Los Alamitos, California, USA, pp. 179-185, Sept. 2004

12 Costilla C, Palacios J, Rodríguez M, Fernández R, Cremades J, Calleja A, *Web Digital Archives Integrated Architecture*, 5th Int. Conf. **Internet Computing** (**IC 2004**), Session: **Web Mining**, Int. MultiConf. **Computer Science& Computer Engineering**, Proc.of **IC 2004**, Arabnia H, Droegehorn O (eds.), ISBN:1932415-44-0, V.1, **CSREA 2004,** Las Vegas, USA, pp.128-134, June 2004

13 Costilla C, Rodríguez M, Palacios J, Cremades J, Calleja A, Fernández R, *A Contribution to Web Digital Archive Integration from the Parliamentary Management System 'SIAP'*, Sixth Int. Baltic Conf. **Data Bases and Information Systems**

(**DB&IS'2004**), Barzdins J. (ed.), ISBN:9984-770-11-7, **Latvijas Univ. Raksti,** Riga, Latvia, pp.481-496, June 2004

14 Costilla C, Cremades J, (invited conference), *e-government: Archivos Digitales en la Web Semántica a partir del Sistema de Gestión Parlamentario 'SIAP'*, in ***e-Gallaecia 2004***, IV Semana Int, de las **TIC**, Santiago de Compostela, pp. 212-220, May 2004

15 Costilla C, Cremades J, Calleja A, Fernández R, Palacios J, *Integración de Archivos Digitales en la Web a partir del Sistema de Gestión Parlamentario 'SIAP'*, IX Cong. Nac. de **Internet, Telecomunicaciones y Movilidad, Internet 2004,** Dep. Legal: M-5613-2004, Madrid, Spain, pp.41-57, Feb. 2004

16 Cavero J, Costilla C, Marcos E, Piattini M and Sánchez A, A *Multidimensional Data Warehouse Development Methodology, Chapter 10*, **Managing Data Mining Technologies in Organizations: Techniques and Applications**' (288 pages), ISBN: 1-59140-057-0 (hardcover), **Parag Pendharkar**, **Idea Group Publishing,** USA, pp. 188-201, 2003

17 Eibe S, Costilla C, Menasalvas E, Acuna C., *DAWIS: Una Arquitectura de Integración Web para el Acceso Integrado a Archivos Digitales*, **VIII J. Ingeniería del Software y Bases de Datos,** ISBN: 84-688-3836-5, Alicante, pp. 583-591, Nov. 2003

18 Sáenz J, Costilla C, Marcos E, Cavero J., *Una Representación en UML del Metamodelo Estándar ISAD(G) e ISAAR(CPF) para la Descripción de Archivos Digitales*, **VIII J. Ingeniería del Software y Bases de Datos,** ISBN: 84-688-3836-5, Alicante, Spain, pp. 519-528, Nov. 2003

19 Costilla C, Calleja A, Cremades J, *Sistema Integrado de Gestión Parlamentaria*: 'SIAP',**Revista Círculo de Usuarios de Oracle, CUORE**, Sección 'Vivat Academia', **Oracle Ibérica**, **Coure,** Madrid, Spain, pp.1-8, Octubre 2003

20 Cavero J, Costilla C, Marcos E, Vela B., *Extending UML for Modelling Queries to Object-Relational Databases*, Fifth Int. Conf. on Enterprise Information Systems (**ICEIS2003**), Vol. 3, ISBN:972-98816-1-8, **Escola Superior de Tecnología de Setúbal,** Angers, France, pp. 535-539, April 2003

21 Costilla C, Calleja A, Cremades J, *Sistema de Información para Ayuntamientos y* Parlamentos: 'SIAP', Boletín de Política Informática, Año XXVI, Nº 6, **Instituto Nac. de Estadística e Informática**, ISSN: 0186-0461, **México, México**, pp. 82-95, 2003

22 Cavero J, Costilla C, Marcos E, Vela B., *Modelado de consultas a BDOR con UML*, **VII J. Ingeniería del Software y Bases de Datos,** ISBN: 84-688-0206-9, El Escorial, pp. 141-150, Nov. 2002

23 Costilla C, Eibe S, Menasalvas E, Sáenz J, Marcos E, Cavero J, Vela B., *DAWIS: Enfoques Preliminares sobre la Arquitectura de Referencia para la Integración de Archivos Digitales en Web*, **JISBD'02**, W. **Red de Excelencia de Bases de Datos en España (RedBD)**, El Escorial, pp. 1-5, Nov. 2002

24 Marcos E, Cáceres P, Cavero J, Vela B, Costilla C, Eibe S, Menansalvas E, Sáenz J, *DAWIS: Sistematización del Desarrollo de Portales para el Acceso Integrado a Archivos Digitales en la Web*, **JISBD'02**, W. **Red de Excelencia de Bases de Datos en España (RedBD)**, El Escorial, pp. 11-16, Nov. 2002

25 Cavero J, Costilla C, Marcos E, Vela B, *Modelado de consultas a BDOR con UML*, VII J. Ingeniería del Software y Bases de Datos, **JISBD'02**, pp. 141-150, El Escorial, Madrid, Nov. 2002.

3 Perspectiva investigadora y Lecciones Aprendidas

Estamos ampliando el alcance de esta investigación, y empezamos a aplicar esta arquitectura para dar servicios a otros dominios cualesquiera. Uno de ellos, y de gran alcance, donde ya investigamos es **e-government** (Ayuntamientos, Administraciones, otras Instituciones, Parlamentos, Archivos, Iniciativas Legislativas, Registros, etc.).

Hemos contribuido en los siguientes foros:

a) Desde el pasado Septiembre, colaboramos en **The European Science Foundation** (ESF) que es responsable del programa ***Towards Electronic Democracy*** (TED), cuyo objetivo global se centra en la e-Democracy (http://www.inf.unibz.it/tcgov2005). Los congresos TCGOV'05 de TED son de muy alto nivel. Sus proceedings se publican por Springer en la Subserie LNAI de la Serie LNCS. El pasado mes de Marzo hemos asistido al último de estos, celebrado en Bolzano (Italia).

a) Colaboramos con la Universidad Inglesa "**School of Computing and Engineering"** de Huddersfield (UK). Como primeros resultados, podemos citar:

- la **publicación conjunta de un libro** y
- la organización de IEEE ADO'05 en **The 3rd Int. Conf. on Information Technology and Applications**. Topic 9: **Agents, Datamining and Ontologies (ADO'05),** publicados por IEEE Computer Society.

b) Cabe destacar **la actual realización de seis tesis doctorales** en estas líneas de investigación.

c) Nuestro trabajo ha dado lugar a colaboraciones con otros grupos de investigación. Hemos contactado con ellos para definir estrechas colaboraciones en futuros proyectos de I+D. Esto se ha materializado en tres casos:

d) Solicitud de un TIN'05 coordinado entre 5 Universidades Españolas. El equipo UPM, además del grupo SINBAD, se ha formado con profesores investigadores de la Universidad Carlos III de Madrid (técnicos y documentalistas).

e) Solicitud de dos proyectos europeos IST, e-Government (uno es IP, el otro STREP). El equipo UPM, además del grupo SINBAD, está formado por profesores de la Universidad Carlos III de Madrid (técnicos, documentalistas y jurídicos) y por profesores de la Universidad de Málaga (técnicos).

Finalmente, hemos sido co-organizadores, revisores o miembros de Comités de Programa en los siguientes Congresos Internacionales:

Co-chair in The IEEE 3rd Int. Conf. on Information Technology and Applications, Topic **Agents, Datamining and Ontologies (ADO'05)**. Hemos sido siete los revisores del grupo SINBAD. http://attend.it.uts.edu.au/icita05/

PC Members and Reviewers in **10th IEEE Symposium on Computers and Communications**. Hemos sido seis los revisores de SINBAD. http://www.comsoc.org/iscc/2005/, http://edas.info/R.cgi?r=285513

PC Members and Reviewers in **Web Semantics 2005 in Database and Expert System Applications, DEXA 2005**. Hemos sido seis los revisores de SINBAD. http://www.dexa.org,

Agradecimientos

Nuestro reconocimiento a la ayuda recibida para investigar. En http://sinbad.dit.upm.es se detalla cada trabajo investigador realizado. Especialmente por la reciente ayuda de TIC02-04050-02-2 (2002-2005) y EDAD-UPM financiado por la Comunidad de Madrid (2003-2004).

RESUMEN

TITULO DEL PROYECTO: "***DiaCrón*** un sistema informático polivalente para su aplicación en la investigación de la arqueología prehistórica"

INVESTIGADORA RESPONSABLE: Carmen- Rosa OLÀRIA PUYOLES

PALABRAS CLAVE: Informática, Arqueología, Prehistoria, Metodología, Excavación, Tipología, Tecnología de la información

El nombre DiaCrón intenta reflejar un sistema de registro informático válido para todo tipo de excavación arqueológica, y muy especialmente para su aplicación en arqueología prehistórica y en los materiales y documentación que se generen en este tipo de estudios.

Diacrón es, básicamente, un Sistema de Gestión de Bases de Datos Relacionables (SGBDR) altamente personalizado, en la que se ha primado la creación de un interfaz amigable y coherente con la metodología de investigación prehistórica, puesto que se entiende que va dirigido a usuarios cuya formación en Tecnologías de la Información no tiene porque ser excesiva.

El objetivo final de este proyecto es la creación de un Software de Registro, Gestión y de apoyo a la interpretación para intervenciones prehistóricas, que sea totalmente funcional, y que será de distribución libre a modo de *freeware*. Para ello, se cuenta con las correspondientes licencias *runtime* que van a permitir la libre difusión del sistema. Además, está previsto la incorporación de las bases de datos originales para favorecer la personalización y adaptación del programa a necesidades más concretas.

El sistema pretende obtener el máximo de distribución posible, y es por ello que se ha elegido un *software* multiplataforma que puede ser ejecutado tanto en plataformas Windows, Mac Os y (bajo *Wine*) en Linux. Además, el programa será también multi-lenguaje, de manera que inicialmente estará disponible en castellano, catalán, inglés y francés, y está abierta la traducción a cualquier otro idioma.

La base metodológica implicada en el programa ha supuesto un importante esfuerzo de recopilación, sistematización y análisis crítico de cómo se efectúan las intervenciones arqueológicas, especialmente en prehistoria. Así, DiaCrón está modelizado a partir de una serie de entidades principales (yacimientos, intervenciones, fases y unidades) alrededor de las cuales se organiza toda la información pertinente referida a estas entidades: los materiales arqueológicos (cerámica, industria lítica, industria ósea, otros materiales, fauna, restos humanos, muestras, etc). También permite almacenar y gestionar todo tipo de información gráfica multimedia (fotos digitales, videos digitales, archivos de sonido, panorámicas, archivos cad, etc).

Para favorecer además la interpretación del conjunto de información almacenada, además, se han añadido toda una serie de funcionalidades y hojas de cálculo que permitirán, entre otras operaciones, tests estadísticos

complejos, aplicables a los conjuntos cerámicos y líticos procedentes de las diferentes intervenciones arqueológicas.

El software irá acompañado, además, de la publicación de un manual con toda la discusión metodológica implicada y la explicación del funcionamiento detallado del programa.

Así mismo se elaborará una página Web alojada en los servidores de la Universitat Jaume I, que va a permitir descargar el sistema, los manuales, consultar las FAQ's y, si procede, las correspondientes actualizaciones. Además también se está analizando la posibilidad de lanzar y mantener un foro público donde los usuarios puedan debatir sus experiencias y dudas, indicar errores o contradicciones en el sistema, y que, al mismo tiempo, fomente el debate sobre el registro y gestión de intervenciones arqueológicas en general y de DiaCrón en particular, así como proponer mejoras en el sistema.

En el estado actual DiaCrón se encuentra en una fase avanzada de su desarrollo. Tras lanzar la beta y distribuirla controladamente entre usuarios cercanos, ha sido ya usado con éxito para el registro de diferentes excavaciones, tanto prehistóricas como protohistóricas e históricas, y las conclusiones extraídas por estos *beta-tester*, han resultado altamente beneficiosas para la corrección y desarrollo definitivo del programa.

Actualmente, próxima ya la finalización del proyecto, se está trabajando en tres direcciones. En primer lugar se está efectuando la redacción del manual, aspecto este que será concluido con toda probabilidad a finales de septiembre de 2005.

En segundo lugar se están acelerando los trabajos de traducción, que se esperan finalizar al mismo tiempo.

Y en tercer lugar, se están realizando los tests estadísticos para la industria lítica, y su integración lo más transparente posible en el sistema DiaCrón, aspecto este que estará concretado antes de que finalice el año en curso.

Análisis y Explotación del Conocimiento Espacio-Temporal en una Web Semántica. Aplicación a la investigación Arqueológica.

(TIC2002-04586-C04-02)

A. Polo, J.M. Fernández, L.J. Arévalo, E. Cerrillo, J.C. Manzano, M. Salas
Departamento de Informática. Universidad de Extremadura
Escuela Politécnica. Avda de la Universidad s/n. Campus Universitario.
10071 – Cáceres (Cáceres)
Teléfono: 927 25 72 49 Fax: 927257202
E-mail: polo@unex.es

Abstract*. In this project our group has been mainly working in two tasks. On the one hand the problem of managing versions in XML documents and, on the other hand, the analysis and the implementation of a mechanism of indexation which could be adapted to the ontology instances. The first one has been solved with two techniques called metamarkup and versioning element that allow us the managing of versions of any document based on the specification XML. The easy management of multiple versioning and the possibility to query the documents by means of standard XML (XQuery and XPath) are some of the advantages of the proposed system. The second work has been the development of an ontology that represents a multidimensional space based on the context that allows us to retrieve knowledge in an efficient way, applying it to the archaeology.*

Introducción

El objetivo de este proyecto es abordar la gestión, consulta y explotación del conocimiento procedente de un conjunto independiente de fuentes de información arqueológicas y museísticas. Para ello, se propone una infraestructura tecnológica que combina técnicas de Bases de Datos y Extracción de Información y que está orientada a un gran volumen de conocimiento distribuido entre las fuentes de información. En concreto, el trabajo desempeñado por nuestro grupo dentro del proyecto se ha centrado en dos partes:

1. *Versionado de documentos XML*.
   Las soluciones actuales, basadas en operaciones deltas XML o mediante la incorporación de aspectos temporales, se caracterizan por no contemplar versionado no lineal y su aplicabilidad está restringida a un determinado lenguaje de marcado, resultando su adaptación a otro bastante complejo. En este proyecto se ha desarrollado un sistema para la gestión de versiones de cualquier documento basado en la especificación XML. Para ello se han propuesto dos técnicas: *elemento de versionado* y *metamarcado* y las primitivas básicas de cambio de cualquier documento de marcado XML, lo que nos permite versionar cualquier otro lenguaje de marcas XML a partir de ellas. Como principales ventajas aporta el soporte ramificado, la consulta mediante estándares XPath y XQuery y su aplicabilidad a cualquier lenguaje de marcado.

2. *Modelo multidimensional de contextos.*
   En la Web semántica nos encontramos con gran variedad de ontologías que describen una parte del conocimiento de un dominio, en nuestro caso ese dominio es la arqueología. Así pues, hemos diseñado una ontología que modela un espacio multidimensional que nos permite definir contextos para ordenar y alinear el conocimiento (ontologías) y facilitar su recuperación.

Versionado de un documento XML

Para su consecución en primer lugar se ha definido una técnica para representar para cada ente o elemento del documento sus versiones asociadas, a partir de los conceptos de bases de datos temporales [1]. Por un lado se ha añadido al documento el *árbol de versionado* [2,3] que representa fielmente qué versiones hay definidas en dicho documento. Una vez almacenada sus versiones, debemos determinar la validez de los elementos del documento asociando a cada uno de ellos un *elemento de versionado* [2,3]. Este se encuentra definido como la unión de un conjunto finito de intervalos de tiempo, que consiste en un conjunto de pares [v1,v2], donde v1 indica el tiempo válido de inicio y v2 el tiempo final. En nuestro caso, y debido a que se ha usado el concepto de versiones en vez de tiempo para representar las distintas versiones, este intervalo representa un camino en el árbol de versionado. Esta particularidad nos permite por un lado tener soporte ramificado y por otro realizar consultas basadas tanto en relaciones temporales de tipo lineal como en relaciones genéricas de versionado no lineal.

La integración de múltiple marcado o versiones en un único documento puede provocar problemas de anidamiento, imposibilitando el tratamiento y la consulta del documento. Para solventar este problema se ha definido la técnica del *metamarcado* [2,3]. Ésta

consiste en transformar todo documento XML en otro documento XML equivalente, sustituyendo cada una de las etiquetas de apertura y cierre del documento original por una marca que la representa denominada metamarca. De este modo, existe una función M que se encarga de transformar un documento XML a un documento metamarcado e igualmente a partir del documento metamarcado, se puede obtener el fichero original mediante una función inversa M^{-1}.

Para comprobar ambas técnicas (elementos de versionado y metamarcado) se experimentó en primer lugar sobre un conjunto significativo de operaciones de cambio sobre documentos XML-Schema [2]. Posteriormente se generalizó esta técnica, no sólo a XML-Schemas, sino a cualquier fichero XML [3]. Para ello se han analizado cuales son las primitivas comunes de cambio de cualquier documento basado en la especificación XML. La principal virtud de nuestro sistema radica en que estas primitivas pueden usarse para versionar cualquier lenguaje de marcado basado en la especificación XML, para lo cual se debe determinar cuáles son las operaciones de cambio del lenguaje y a continuación describir cada una de estas primitivas en base a las operaciones básicas XML.

Modelado multidimensional de contextos

El contexto son las circunstancias, objetos y condiciones que subyacen explícita o implícitamente al usar varias ontologías sobre un mismo dominio. En nuestro caso el contexto debe reflejar aquellos aspectos básicos de la realidad a que se refieren de forma común las ontologías descritas en nuestro sistema.

Una de las mejores formas de describir el contexto es utilizar una ontología, llamada Ontología de Contexto OC [4]. La ontología de contexto se define como un conjunto de conceptos que pueden tener relación con las ontologías almacenadas pero que no se reflejan en éstas por no ser relevantes para el sistema, pero sí para la comprensión de éstos.

La utilización de una ontología nos aporta una serie de ventajas como son la utilización de un vocabulario compartido y la posibilidad de razonamiento mediante las relaciones entre conceptos. Se han utilizado una serie de conceptos genéricos, fácilmente compresibles y adaptables a cualquier situación: Espacio, Tiempo, Materia, Agente y Acción. Pensamos que con estos cinco conceptos podemos representar la mayoría de situaciones del mundo; esto es debido a la equivalencia que existe entre la estructura del lenguaje natural y nuestra ontología [5].

Por último, definimos el espacio multidimensional como un espacio de n dimensiones siendo n igual al número de conceptos definidos en la ontología OC. Cada dimensión estará asociada a un concepto de OC y sus valores serán las instancias de ese concepto. De esta manera los contextos se describen como un conjunto de instancias de OC que forman un subespacio que contiene las ontologías descritas bajo ese contexto.

Finalmente se ha desarrollado un prototipo del sistema aplicado al dominio de la arqueología llamado ArqueOnto [5,6]. En su desarrollo se ha experimentado con dos ontologías para la arqueología diseñadas dentro del proyecto por dos grupos distintos y con diferentes metodologías.

Conclusiones

Con la Web semántica nos encontramos que la información de la Web empieza a estar codificada mediante ontologías, y más en concreto con documentos XML y surge la necesidad de versionar esta información. Los sistemas actuales de versionado sobre documentos de marcado XML tienen la principal desventaja de que únicamente son aplicados sobre un lenguaje concreto y no tienen soporte para versionado ramificado. En este proyecto se ha presentado un sistema de versionado operacional de estado para la gestión de versiones de documentos de marcado XML independientemente de su especificación.

La idea de estructurar el conocimiento en un espacio multidimensional basado en el contexto nos proporciona otro mecanismo de acceso a distintas versiones del conocimiento en función del contexto.

Referencias

[1] Arévalo L., Polo A., Salas M, Manzano J. M2: Una técnica para la integración de versiones de XML Schemas. VIII JISBD. Alicante. 2003.

[2] Arévalo L., Polo A., Salas M, Manzano J. Sistema de versionado genérico en XML. IX JISBD. Málaga. Noviembre 2004.

[3] Arévalo L., Polo A., Salas M, Manzano J. Múltiple Markups in XML Documents. ICWE '03. Oviedo. 2003. LNCS nº 2722.

[4] Fernández J., Polo A., Cerrillo E. Bases for the creation of an ontology in the context of Archaeology. CAA2005, Tomar (Portugal)

[5] Cerrillo E., Fernández J., Polo A., Ontologías: un nuevo método de explotación y difusión del conocimiento arqueológico. Aplicaciones preliminares al yacimiento de Los Barruecos (Cáceres)". II IAIGA, Córdoba (España), 2005.

[6] Fernández J., Polo A., M-SW: Meta-Servicios Web mediante Oracle-XDB. Aplicación a sistemas de gestión de Museos con autentificación multilingüe. XIV Congreso Nacional usuarios de Oracle, Zaragoza. 2004.

CRISOL: Generación automática de instancias ontológicas desde fuentes de datos semi-estructuradas

R. Berlanga, M. J. Aramburu, D. M. Llidó, I. Sanz, J. M. Pérez, R. Danger, J. Paraire, E. Jiménez

Departamento de Lenguajes y Sistemas Informáticos. Universitat Jaume I.
ESTCE. Campus Riu Sec.
12071 – Castellón (España)
Teléfono: 964 72 8367 Fax: 964 72 8435
E-mail: berlanga@uji.es

Abstract*. En este artículo presentamos los objetivos planteados inicialmente en el sub-proyecto CRISOL (TIC2002-04586-C04-03), así como los principales resultados obtenidos. Este se plantea como una primera aproximación al análisis semántico de recursos extraídos desde diversas fuentes de información sobre investigación arqueológica. Para ello, se define una ontología de referencia, y se proponen herramientas para la obtención de los objetos semánticos asociados, su indexación y posterior consulta y análisis.*

1 Introducción

En este proyecto se aborda la problemática de gestionar, consultar y explotar el conocimiento procedente de un conjunto independiente de fuentes de información arqueológicas y museísticas. Para ello, se propone una infraestructura tecnológica común que combina técnicas de *Bases de Datos* y *Extracción de la Información*, y que está orientada al tratamiento eficiente del conocimiento distribuido en diversas fuentes de información. El resultado final del proyecto pretende establecer la base tecnológica de lo que será una Web Semántica para las investigaciones Arqueológicas y de los Museos asociados.

La infraestructura propuesta se construirá a partir de dos elementos básicos: 1) un prototipo de una arquitectura para el procesamiento de consultas basado en ontologías, que se extenderá posteriormente con lo resultados obtenidos de la investigación desarrollada, y 2) un conjunto de *ontologías* que describan la semántica de cada fuente de información y sus dimensiones temporales y espaciales, las cuales son esenciales para la investigación arqueológica y museística.

Sobre esta infraestructura se diseñarán una serie de *servicios finales avanzados* orientados a la explotación del conocimiento. Estos servicios aportarán una visión conceptual común a los investigadores de Arqueología, y permitirán que éstos puedan compartir, analizar y explotar los resultados de sus investigaciones.

2 Desarrollo del proyecto

Las tareas a cargo del grupo de investigación de la Universitat Jaume I (UJI) se centran en el estudio del dominio de la aplicación, la generación automática de instancias ontológicas, y su posterior análisis.

2.1. Estudio del dominio de la aplicación

En esta tarea se ha realizado un estudio detallado del dominio de la aplicación final, de sus fuentes de información y de las ontologías necesarias para describir su conocimiento. Actualmente, no existe ninguna ontología específica sobre Arqueología, y solo se dispone de ontologías generales y esquemas XML para la descripción de patrimonio cultural, como por ejemplo *Conceptual Resource Model* (CRM) y XSTAR. Estos son claramente insuficientes para nuestros propósitos, ya que son demasiado genéricos y no contienen la terminología específica requerida en nuestro dominio. Por esta razón, durante la primera fase del proyecto se definió una ontología para el ámbito de la Arqueología Prehistórica, especialmente orientada a las culturas locales.

Los conceptos y relaciones de esta ontología fueron definidos por el grupo de Arqueología Prehistórica de la UJI, basándose principalmente en la herramienta DiaCron que se desarrolla en su sub-proyecto [1]. Sobre la ontología se han definido tres dimensiones que recogen los distintos aspectos de la investigación arqueológica:

- La descripción de los hallazgos como Patrimonio Cultural, lo cual se logra ligando los conceptos de la ontología con los generales del CRM.
- La descripción de los métodos utilizados por los investigadores, tanto en los procedimientos de excavación como en la clasificación de los materiales hallados. Debido a las discrepancias que existen sobre los métodos, esta dimensión puede verse como distintas versiones de la ontología común [2].
- La dimensión interpretativa de los hallazgos, principalmente la que asocia una funcionalidad y

una datación relativa a cada yacimiento. Esta parte de la ontología contempla la parte de conocimiento a intercambiar con otros investigadores. Por esta razón, la dimensión interpretativa debe ser común a todos ellos.

2.2. Generación de Instancias Ontológicas

A lo largo de esta actividad se desarrollaron las componentes necesarias para la generación de instancias ontológicas que alimentarán finalmente el índice del procesador de consultas [3]. La mayor parte de la información disponible para poblar la ontología propuesta se encuentra en las denominadas Memorias de Excavación, que son documentos de texto poco estructurados. Existen otras fuentes de datos bien estructuradas, como son las Cartas Arqueológicas, pero éstas suelen proporcionar pocos detalles sobre los yacimientos. Todo ello implica un esfuerzo considerable de estructuración y extracción de la información para la creación de las instancias de la ontología.

En una primera aproximación se estudió la aplicación de técnicas de minado de texto para la obtención de patrones frecuentes que pudiesen asociarse directamente a instancias de la ontología [4]. Aunque esta aproximación obtiene una buena precisión, deja fuera mucha información interesante para las instancias que no aparece con frecuencia. Con el fin de obtener instancias más completas, se ha diseñado un algoritmo que trata de inferir las instancias a partir de las palabras y entidades reconocidas en segmentos de texto [5]. Para ello se utilizan las relaciones semánticas definidas en la ontología, tales como la generalización, agregación y la cardinalidad de las relaciones. La principal ventaja de esta aproximación es que no requiere ningún análisis sintáctico de las sentencias, lo cual lo hace muy eficiente.

Por otro lado, se han definido algoritmos de agrupamiento automático de documentos, para crear grupos semánticos de acuerdo con la ontología [6]. Estos algoritmos tienen en cuenta las dimensiones espacio-temporales, los conceptos extraídos de los textos, y la estructura de los documentos [7].

2.3. Minado de instancias ontológicas

Para el análisis y explotación de las instancias generadas, en el proyecto se proponen dos técnicas complementarias. La primea de ellas se basa en el minado de las instancias para la obtención de patrones interesantes. Para ello, se ha diseñado un algoritmo de minado para objetos heterogéneos [8]. Se trata de una versión preliminar, que requiere un estudio más profundo debido a su alta complejidad computacional. La segunda de ellas consiste en un modelo multidimensional que combina la información estructurada de las instancias con la información textual de los documentos [9]. Así, el analista puede expresar consultas en las que intervengan las variables de análisis deseadas y un conjunto de palabras clave relacionadas. El modelo recupera entonces todas las instancias de interés y las puntúa según su relevancia a la consulta. El resultado final es un cubo multidimensional donde los hechos tienen asociados un índice de relevancia, y donde se pueden aplicar tanto las técnicas tradicionales OLAP, como los algoritmos de minado anteriores.

3 Conclusiones

El proyecto CRISOL plantea una primera aproximación a la problemática de generar y analizar instancias ontológicas obtenidas a partir de textos y datos (semi) estructurados. Los resultados finales han sido satisfactorios, especialmente en la definición de la ontología y la obtención de sus instancias. Las tareas de análisis planteadas son muy novedosas, a la vez que desafiantes por su complejidad, y requieren por tanto un estudio más profundo en el futuro.

Agradecimientos

Este trabajo ha sido financiado por el proyecto CICYT TIC2002-04586-C04-03.

Referencias

[1] Proyecto DIACRON. TIC2002-04586-C04-04.

[2] Proyecto UEX. TIC2002-04586-C04-02.

[3] Proyecto UMA. TIC2002-04586-C04-01.

[4] Danger, R., Berlanga, R., Ruíz-Shulcloper, J.: CRISOL: An approach for automatically populating a Semantic Web from Unstructured Text Collections. *Lecture Notes in Computer Science 3180*, 2004.

[5] Danger, R., Sanz, I. Berlanga, R., Ruíz-Shulcloper, J.: A proposal for the automatic generation of instances from unstructured text. *Lecture Notes in Computer Sicence 3287*, 2004.

[6] Pons, A., Berlanga, R., Ruíz-Shulcloper, J., Pérez, J. M.: Jerartop: A new Topic Detection System. *Lecture Notes in Computer Science 3287*, 2004.

[7] Sanz, I.; Pérez, J.M.; Berlanga, R.; Aramburu, M.J : XML Schemata Inference and Evolution. *Lecture Notes on Computer Science 2736*, 2003.

[8] Danger, R., Ruíz-Shulcloper, M.J., Berlanga, R.: ObjectMiner: A New Approach for Mining Complex Objects. *Proc. of ICEIS*, 2004.

[9] Pérez, J. M., Berlanga, R., Aramburu, M.J.: A Document Model based on Relevance Modelling Techniques for Semi-Structured Information Warehouses. *Lecture Notes in Computer Science 3180*, 2004.

CRISOL: Una Plataforma Básica para la Evaluación de Consultas, Mediación e Interoperabilidad en la Web Semántica

Jose F. Aldana, A. Cesar Gómez, Ismael Navas, Mª del Mar Roldán

Departamento de Lenguajes y Ciencias de la Computación. E.T.S. de Ingeniería Informática.
Universidad de Málaga. 29071 Málaga (España).
E-mail: jfam@lcc.uma.es

Abstract. *Se presentan los objetivos del proyecto TIC2002-04586-C04-04 dentro del proyecto CRISOL. Estos son: el desarrollo de una plataforma básica para dar soporte al procesamiento de consultas y la mediación y la interoperabilidad en la Web Semántica. Se describen la arquitectura de la plataforma desarrollada, los resultados obtenidos, las líneas de aplicación práctica y las nuevas líneas de investigación abiertas.*

1 Introducción

Este subproyecto se plantea el objetivo de desarrollar una plataforma básica para dar soporte al procesamiento de consultas, la mediación y la interoperabilidad en la Web Semántica.
Sobre esta plataforma se desarrollarán las aplicaciones de los otros subproyectos de CRISOL y dentro de ella se desarrollaran diversas líneas de investigación básica de este subproyecto: 1) Evaluación y Optimización de Consultas distribuidas; 2) Diseño de Bases de Conocimiento en el Entorno de la Web Semántica; 3) Almacenamiento y Recuperación Eficiente del Conocimiento y; 4) Infraestructura Para la Mediación Semántica.

2 Plataforma

Se seleccionó OWL como lenguaje para describir ontologías y se diseñó una arquitectura que utilizaba la herramienta Protege (ver figura 1) como el núcleo con la finalidad de minimizar los costes de desarrollo y aprovechar su extensibilidad gracias a su capacidad de incorporar diversos tipos de plugins (por ejemplo, un motor de inferencias y una interfaz de consultas para el mismo). Estos plugins están disponibles en el servidor web del proyecto: www.khaos.uma.es/AECETWS.

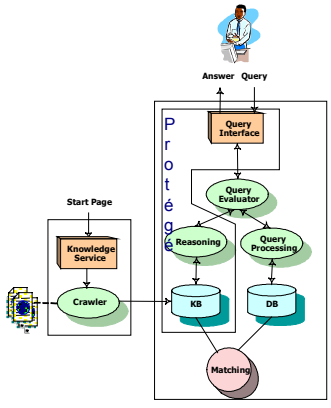


Figura 1. Componentes de la Arquitectura Propuesta

Este prototipo ofrece toda la funcionalidad pretendida para la aplicación final. Y partiendo de él se van a incluir mejoras progresivas: almacenamiento persistente; mecanismos de evaluación de consultas y razonamientos, basados en las lógicas de descripciones; una interfaz gráfica de consulta para lógicas de descripciones; y un crawler que inserte conocimiento extraído de documentos anotados semánticamente en la base de conocimiento Estas extensiones que se muestran en la figura 2 se describen en la sección 3.

3 Extensiones y Trabajos en Curso

La plataforma desarrollada da cabida a varias líneas y trabajos de investigación en torno a un tema central: *la evaluación de consultas y la recuperación eficiente del conocimiento en la Web Semántica*.

3.1 Evaluación y Optimización de Consultas distribuidas

Actualmente nos encontramos desarrollando técnicas para este nuevo entorno. Para ello nos basamos en tres aspectos fundamentales: 1) enriquecimiento del lenguaje aumentando la interpretación semántica de las estructuras de datos; 2) preservación de la eficiencia y; 3) distribución de la optimización/ evaluación.

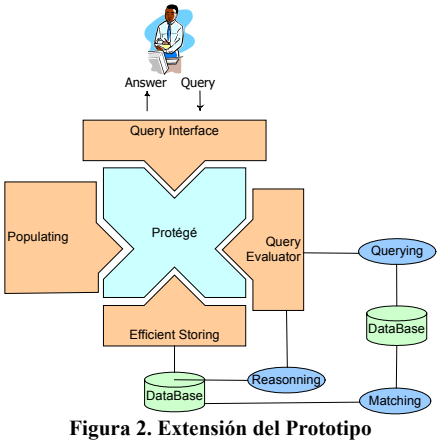


Figura 2. Extensión del Prototipo

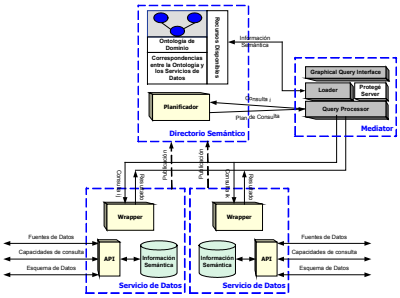


Figura 3. Infraestructura para la Integración Dinámica de Fuentes de Datos

Aunque Protégé posee su propio sistema de evaluación de consultas, para poner en práctica estas técnicas dentro del marco de trabajo del proyecto se sustituira el mismo por otro más completo.

3.2 Diseño de Bases de Conocimiento en el Entorno de la Web Semántica

Para que los sistemas interoperables puedan comunicarse, debe existir un acuerdo sobre la codificación de los datos así como sobre el significado de los mismos. Esto puede conseguirse haciendo un diseño de la base de datos controlado por ontologías. Utilizando ontologías como punto de partida en el diseño de esquemas de bases de datos para aplicaciones específicas. Se han definido nuevas medidas de similitud así como modelos de representación y razonamiento sobre mappings que puedan mejorar los procesos de modelado conceptual así como el procesamiento y/o la optimización semántica de consultas.

3.3 Almacenamiento y Recuperación Eficiente del Conocimiento

La investigación en este campo consiste en la definición de una metodología para representar físicamente una ontología, teniendo en cuenta la estructura de la ontología, los patrones de consulta y la aplicación que le vamos a dar a la ontología, para aportar escalabilidad y eficiencia a la web semántica. Esta representación no es única, sino que es un compromiso que depende de la expresividad del lenguaje de consultas y de la expresividad requerida por el dominio de aplicación (las capacidades de razonamiento necesarias).

3.4 Infraestructura para la Mediación Semántica

Se propone una arquitectura para la mediación semántica que permite la integración dinámica de fuentes de datos. Presenta mejoras sustanciales frente a la mediación tradicional. Ofrece un alto grado de desacoplamiento entre wrappers y mediadores y la posibilidad de reutilizar aquellos y proporciona los elementos necesarios para obtener una mayor interoperabilidad y se abre la puerta a la integración dinámica. Ver figura 3.

4 Indicios de Calidad

Dentro de este proyecto se han publicado los siguientes artículos: 6 en revistas internacionales [1] [2] [3] [4] [5] [6]; 1 en revistas internacionales [7]; 3 en LNCS [8] [9] [10]; 13 en conferencias internacionales [11] [12] [13] [14] [15] [16] [17] [18] [19] [20] [21] [22] [23] y, 10 en conferencias nacionales.

Agradecimientos

Este trabajo ha sido financiado por el proyecto CICyT TIC-2002-04586-C04-04.

Referencias

[1] M.M. Roldan, I. Navas, J. F. Aldana. “Using Knowledge for Enhancing Queries in the Semantic Web”. Proceedings of 4th International Conference on Knowledge Management. I-KNOW 2004. Journal of Universal Computer Science (J.UCS). Graz. Austria, 2004. ISBN: 0948-695x. ISSN: 0948-6968. Pags. 345-352.

[2] N. Moreno, I. Navas, A.C. Gomez, M.M. Roldan, J. F. Aldana. “Musa-K: A Practical Step to Integrate Databases and Semantic Web Technologies”. Proceedings of 4th International Conference on Knowledge Management. I-KNOW 2004. Journal of Universal Computer Science (J.UCS). Graz. Austria, 2004. ISBN: 0948-695x. ISSN: 0948-6968. Pags. 388-396.

[3] N.Moreno, I. Navas, J.F. Aldana. “Putting the Semantic Web to Work with DB Technology”. IEEE Bulletin of the Technical Committee on Data Engineering. Diciembre 2003. Pags. 49-54.

[4] José-Francisco Aldana-Montes, Antonio-César Gómez-Lora, Nathalie Moreno-Vergara, and María del Mar Roldán-García. “Interrogare il Semantic Web”. Versión italiana de UPGRADE: the European On-line Magazine for the IT Proffesional., edición electrónica. Agosto, 2002.

[5] J.F. Aldana, A.C. Gómez, N. Moreno, M.M. Roldán. “Querying the Semantic Web: Feasibility Issues”. Journal of CEPIS (Council of European Professional Informatics Societies). Septiembre 2002. ISSN 1684-5285.

[6] J.F. Aldana, A.C. Gómez, N. Moreno, M.M. Roldán. “Viabilidad Práctica de la Evaluación de Consultas en la Web Semántica”. Novática, Agosto 2002. ISSN 0211-2124.

[7] Ismael Navas-Delgado, María del Mar Roldán-García, José Francisco Aldana-Montes. “Deep Crawling in the Semantic Web. In Search of Deep Knowledge”. The Fifth International Conference on Web Information Systems Engineering. Wise 2004. November 22-24, 2004, Brisbane, Australia. Lecture Notes in Computer Science, vol 3306. ISBN: 3-540-23894-8. ISSN: 0302-9743. Pag. 541-546.

[8] Ismael Navas-Delgado, María del Mar Roldán-García, José Francisco Aldana-Montes. “Kreios: Towards Semantic Interoperable Systems”. Third Biennial International Conference on Advances in Information Systems. ADVIS’2004. 20-22 October, 2004. Izmir, Turkey. Lecture Notes in Computer Science, vol 3261. ISBN: 3-540-23478-0. ISSN: 0302-9743. Pags. 161-171.

[9] José Francisco Aldana-Montes, Ismael Navas-Delgado, María del Mar Roldán-García. “Solving Queries over Semantically Integrated Biological Data Sources”. Proceedings of the Fifth International Conference on Web-Age Information Management WAIM 2004. Lecture Notes of Computer Science, vol. 3129. Dalian, China. 15 - 17 July 2004. ISBN: 3-540-22418-1, ISSN: 0302-9743. Pags. 249-258.

[10] María del Mar Roldán-Garcia, Ismael Navas-Delgado, José F. Aldana. “A Design Methodology for Semantic Web Database-Based Systems”. The 3rd International Conference on Information Technology and Applications. Agents, Datamining and Ontologies (ADO’05). July 4-7, 2005, Sydney, Australia. 6 páginas.

[11] Ismael Navas-Delgado, Carlos Rodriguez-Caso, Maria del Mar Roldan-Garcia, Miguel Angel Medina, Jose F. Aldana-Montes. “A Semantic Tool for Analyzing Interaction Among Transcription Factors”. IADIS International Conference, Applied Computing 2005, Algarve, Portugal, 22-25 February 2005. Pags. 512-518.

[12] Aldana, José F., Hidalgo-Conde, Manuel, Navas, Ismael, Roldán, María del Mar, Trelles, Oswaldo. “Bio-Broker: A biological data and services mediator system”. IADIS International Conference, Applied Computing 2005, Algarve, Portugal, 22-25 February 2005. Pags. 527-534.

[13] J.F. Aldana, M.Hidalgo-Conde, I. Navas, M. del Mar Roldán, P.A. de Alarcón and O.Trelles. “An Open Architecture for the Dynamic Integration of Biological Data Sources”. International Conference on Computing, Communications and Control Technologies: CCCT'04. August 14-17, 2004 - Austin, Texas, USA. ISBN: 980-6560-17-5. Pags. 239-244.

[14] Ismael Navas-Delgado, Nathalie Moreno-Vergara, Antonio C. Gómez-Lora, Maria del Mar Roldán-García, Iván Ruiz-Mostazo, José F. Aldana-Montes. “Embedding Semantic Annotations into Dynamic Web Contents”. International Workshop on Web Semantics - WebS 2004 in conjunction with DEXA 2004 14th International Conference on Database and Expert Systems Applications. Zaragoza, Spain. August 30 - September 3, 2004. IEEE Computer Society Press. ISBN: 0-7695-2195-9. ISSN: 1529-4188. Pags. 231-235.

[15] J.F. Aldana, M. Roldán, I. Navas, A.J. Pérez, O. Trelles. “Integrating Biological Data Sources and Data Análisis through Mediators”. Proceedings of the 2004 ACM Symposium on Applied Computing. SAC 2004. Nicosia, Chipre. 2004. ISBN: 1-58113-812-1. (Available online only). 5 páginas.

[16] J. F. Aldana, I. Navas, M. M. Roldan. “Semantic Integration of Digital Libraries”. Proceedings of 6th International Conference on Enterprise Information Systems. ICEIS 2004. Oporto, Portugal. 2004. ISBN: 972-8865-00-7. Pags. 313-318.

[17] I. Navas, J. F. Aldana. “Towards Conceptual Mediation”. Proceedings of 6th International Conference on Enterprise Information Systems. ICEIS 2004. Oporto, Portugal. 2004. ISBN: 972-8865-00-7. Pags. 169-176.

[18] J.F. Aldana, A.C. Gómez, N. Moreno, I. Navas, M.M. Roldán. “Database Techniques for the Semantic Web”. Proceedings of 1ère Conférence en Sciences et Techniques de l’Information et de la Communication (CoPSTIC’03). Rabat, Marruecos. 2003. Pags. 184-188.

[19] J.F. Aldana, A.C. Gómez, N. Moreno, I. Navas, M.M. Roldán. “Database Techniques for the Semantic Web”. Proceedings of 1ère Conférence en Sciences et Techniques de l’Information et de la Communication (CoPSTIC’03). Rabat, Marruecos. 2003.

[20] Natalia Moreno Vergara, Francisco Aldana Montes, Pedro Alarcón, Oswaldo Trelles Salazar. Distance Learning Platform for Bioinformatics. Internacional Conference on Multimedia ICTS in Education, 2003. Badajoz, España.

[21] Francisco Aldana Montes. New Methods for Organization in the Web of Data Concerning Amine-Related Metabolic and Regulatory Pathways. Cost 922 Health Implication of Dietary Amines: 2nd Workshop on Amines and Food Safety. 2003.Málaga, Spain.

[22] J.F. Aldana, M.M. Roldán, A.C. Gómez, N. Moreno, A.J. Nebro. “Metadata Functionality for Semantic Web Integration”. Proceedings of the Seventh International ISKO Conference. 10-13 July 2002. Granada, Spain. ISBN: 3-89913-247-5. ISSN: 0938-5495.

ECIM: Un Entorno Computacional para la Intervención Médica. Desarrollo de la Plataforma Básica e Integración Hospitalaria-Canarias.

Juan Ruiz-Alzola, Miguel Ángel Rodríguez-Florido, Rubén Cárdenes, Eduardo Suarez-Santana
Centro de Tecnología Médica. Dpto. de Señales y Comunicaciones. Universidad de Las Palmas de Gran Canaria
Pabellón B de Telecomunicación. Campus de Tafira, s/n
35017 – Las Palmas de Gran Canaria
http://www.ctm.ulpgc.es/
E-mail: {jruiz, marf, ruben, eduardo}@ctm.ulpgc.es

***Abstract**. The project ECIM: A Computation Environment for Medical Intervention belongs to spanish R&D Plan 2000-2003. From the research point of view, it has led to important contributions in areas of medical image processing such as filtering, registration and segmentation. It has also allowed to develop radiological software to be integrated in real clinical environments. However, its main contribution has been to support the creation of a spanish research group with high international visibility and impact in the field of medical imaging.*

1 Introducción

Este documento expone los resultados principales del proyecto de investigación científica y desarrollo tecnológico de referencia TIC2001-3808-C02-01, que se enmarca dentro del Plan Nacional I+D+i 2000-2003. Nos referiremos a él como ECIM, acrónimo de un Entorno Computacional para la Intervención Médica. Dicho proyecto ha sido ejecutado por el Centro de Tecnología Médica, a partir de ahora CTM.

El objetivo que se ha perseguido en este proyecto, es la elaboración de un sistema computerizado de asistencia a la planificación quirúrgica basado en imágenes tridimensionales e integrable en un entorno clínico real. De forma paralela, se ha planteado su utilidad en aplicaciones diagnósticas, científicas y docentes.

Los resultados de ECIM abarcan tres campos bien diferenciados que se exponen en sendas secciones siguientes: el campo científico-técnico, el campo socio-económico, así como el campo de infraestructura de gestión y desarrollo de la I+D+i.

2 Resultados Científico-Técnicos

2.1 Investigación

Dentro del área de investigación, se ha trabajado intensamente en el procesado de señales multidimensionales, en particular en las áreas de filtrado [1], registrado [2,3] (esta última publicación premiada) y segmentación [4] de imágenes médicas, con respectivas tesis leídas en cada una de estas áreas. Además, se ha contemplado en la elaboración de los algoritmos su eficiencia computacional para su desarrollo en tiempos clínicamente aceptables, haciéndose necesario en algunos casos estrategias de paralelización de los mismos [4].

Entre los resultados clínicos, destaca la elaboración de atlas de cerebro sano, para el posterior estudio de patologías de esclerosis múltiple [5], y de rodilla para el estudio de patologías asociadas al cartílago.

2.2 Desarrollo

Actualmente se ha concluido el desarrollo de un software de aplicación (*véase* Fig. 1), basado en el software de código abierto *Slicer*, de asistencia al diagnóstico y a la intervención, basado en las librerías *vtk* e *itk*, listo para ser insertado en entornos médicos, ya que se ha integrado en él la conectividad al sistema radiológico de archivo de imágenes médicas (PACS).

Se oferta además en el marco del proyecto, un *toolbox* desarrollado sobre *Matlab*, para el procesado de señales multidimensionales, genérico para datos tensoriales.

3 Resultados Socio-Económicos

Entendiendo por innovación el cambio para el beneficio socio-económico, a continuación se detallan los resultados socio-económicos del mismo.

En primer lugar, en cuanto a las personas, el entorno científico-técnico que ha creado ECIM ha formado entre otros a dos becarios que han sido requeridos por el mercado laboral y trabajan ahora fuera de la universidad. Además, recibimos mensualmente solicitudes de estudiantes extranjeros que desean formarse con nosotros.

En cuanto a la tecnología, se ha transferido, en forma de estancia en el extranjero, una parte de la tecnología de segmentación de alto rendimiento desarrollada.

En cuanto al impacto social, el crecimiento que ECIM ha proporcionado al CTM ha interesado a las instituciones públicas locales, con la firma de un convenio entre el Instituto Tecnológico de Canarias y la Universidad de Las Palmas de Gran Canaria, cuyo objetivo es la integración del sector I+D (donde entraría el CTM), el sector clínico y el sector industrial.

4 Innovación en la I+D+i

Aparte de la innovación en la producción científico-técnica, existe un componente importante de innovación en los sistemas, crítico en el rendimiento del proyecto.

Destacamos aquí la implantación de un sistema de información (a partir de código abierto) que atesora el capital intelectual generado, sirviendo a su vez como marco de trabajo (web dinámica, planificador de tareas, bitácora, gestor documental, calendario web, foro de consulta y listas de distribución).

En el área de producción de desarrollo, se han definido también los procesos de elaboración de software radiológico, desde el contacto con el especialista, las herramientas, lenguajes y plataformas de desarrollo, hasta las iteraciones que dan lugar al producto final listo para ser validado.

5 Conclusiones

Las iniciativas del ministerio han satisfecho el objetivo general de la creación en España de un grupo multidisciplinar de referencia internacional en el campo de la radiología computacional. Es remarcable aquí el intercambio de estudiantes de doctorado con instituciones como Harvard Medical School, INRIA y Houston Medical School.

Este documento recoge sólo las aportaciones más importantes del proyecto, pudiendo tener una visión más detallada a partir de la web del mismo: http://www.ctm.ulpgc.es/ecim/

6 Líneas Futuras

Hoy en día, investigamos las posibilidades de las herramientas de procesado en la tecnología de adquisición tridimensional de bajo coste, el ultrasonido 3-D, así como las posibilidades que ofrece la elastografía como técnica de adquisición.

Respecto al desarrollo, estamos estudiando las posibilidades que ofrece el desarrollo orientado a componentes para la reusabilidad del software.

Desde la perspectiva de la innovación, nos quedan aún por establecer los sistemas de acceso a los mercados que permitan una transferencia tecnológica más fluida y el "market-pull", esto es, la orientación a cubrir la demanda.

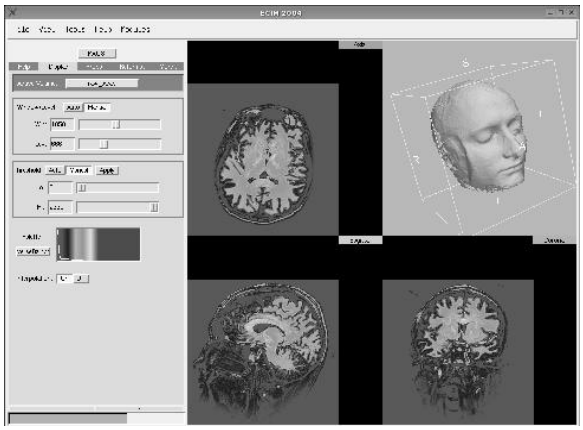


Fig. 1. Software radiológico de aplicación

Agradecimientos

El Centro de Tecnología Médica, en su totalidad, agradece la financiación de este proyecto de investigación y desarrollo tecnológico, que ha potenciado enormemente la creación de un nuevo foco estatal de I+D+i, e igualmente manifiesta su satisfacción por la gestión llevada a cabo por el Departamento de Tecnologías de la Producción y las Comunicaciones del Ministerio de Educación y Ciencia.

Referencias

[1] Miguel Ángel Rodríguez-Florido. "Procesado Anisótropo de Campos Tensoriales Multidimensionales y sus Aplicaciones al Filtrado y Segmentación de Imágenes Médicas". Julio 2003. Universidad de Las Palmas de Gran Canaria.

[2] E. Suarez-Santana. "Un Marco General para la Normalización Geométrica de Datos Tensoriales y su Aplicación al Registrado de Imágenes Médicas". Julio 2003. Universidad de Las Palmas de Gran Canaria.

[3] E. Suarez-Santana, C.F. Westin, J. Ruiz-Alzola. "Nonrigid Registration Using Regularized Matching Weighted by Local Structure". LNCS 2489, pp.581-589.

[4] Rubén Cárdenes. "Esquemas Eficientes de geometría Computacional Aplicados a la segmentación de Imágenes Medicas". Junio 2004. Universidad de Las Palmas de Gran Canaria.

[5] Rubén Cárdenes, Simon K. Warfield, Elsa M. Macía, José Aurelio Santana, Juan Ruiz-Alzola, "An Efficient Algorithm for Multiple Sclerosis Lesion Segmentation from Brain MRI", EUROCAST 2003, pp 542-551.

ECIM: Un Entorno Computacional para la Intervención Médica. Desarrollo de la Plataforma Básica e Integración Hospitalaria. (TIC2001-3808-C02-02)

C. Alberola López
Departamento de Teoría de la Señal, Comunic. e Ing. Telemática. Universidad de Valladolid
ETSI de Telecomunicación. Cam. Cementerio s/s. Campus Miguel Delibes
47011 - Valladolid (Valladolid)
Teléfono: 983 42 36 66 Fax: 983 42 36 67
E-mail: caralb@tel.uva.es

Abstract *We describe a brief overview of the tasks carried out within the scope of grant TIC2001-3808-C02, a coordinated grant with two partners, namely, Universidad de Las Palmas de Gran Canaria (ULPGC) and Universidad de Valladolid (UVA). In this summary we stress the contribution of the UVA partner and in the reference section we include a subset of the journal and conference papers these funds have given rise to.*

1 Introducción

El proyecto TIC2001-3808-C02 prevé la construcción de una herramienta para la intervención médica en la que se incluirán algoritmos de procesado para visualización de datos y ayudas al diagnóstico. Para tal fin, hemos partido de herramientas construidas por el grupo de trabajo en el pasado [1]; la creciente relación de nuestro grupo con el *Surgical Planning Laboratory* (SPL) de la Universidad de Harvard, ha hecho aconsejable hacer uso de una herramienta similar, generada por este grupo y denominada 3D-Slicer, y sobre ella montar los desarrollos generados en el proyecto.

Tales desarrollos se han centrado básicamente en dos frentes, a saber, ámbito telemático y ámbito de procesado. En lo que sigue realizaremos una breve descripción de cada uno, proporcionando múltiples referencias en relación con las publicaciones surgidas del proyecto, y en las que se detallan los algoritmos y desarrollos concretos.

2 Ámbito telemático

La mencionada herramienta 3D-Slicer es una herramienta monousuario; sin embargo hemos observado que los estudiantes de medicina que hacen uso de la misma en el SPL de forma frecuente necesitan resolver dudas, para lo cual acuden a un experto. Entendemos que tal actividad se puede llevar a cabo de forma cómoda mediante una extensión colaborativa del 3D-Slicer, la cual ha desarrollado nuestro grupo, y ha recibido el nombre de *Group-Slicer* [2]. Para llevar a cabo esta extensión colaborativa se ha tratado de respetar al máximo los procedimientos médicos que ejecutan estudiantes y especialistas de forma cotidiana, en particular, el interfaz gráfico de usuario ha permanecido inalterado. Se ha realizado una análisis de prestaciones, en términos de tráfico, y se ha visto que la solución propuesta es perfectamente viable, incluso en redes lentas.

3 Ámbito de procesado

El proyecto contempla varias líneas de procesado, a saber, segmentación, filtrado e interpolación y registrado, siempre con orientación hacia la creación de esquemas de ayuda al diagnóstico. Algunos detalles en relación con estas actividades:

3.1 Segmentación

Las contribuciones en segmentación se han realizado en varios frentes. El primero de ellos es el frente de ecografía renal, en el cual se ha desarrollado un método semiautomático de extracción de contornos renales [3], oportunamente validado [4]. Se ha abordado también el problema de extracción automática de contornos de huesos de manos mediante contornos activos. En esta dirección cabe destacar [5].

3.2 Filtrado e interpolación

Las contribuciones en filtrado e interpolación de imagen para ulteriores procesados son numerosas y en diversas direcciones. Hemos contribuido al filtrado de imagen ecográfica [6, 7], al filtrado mediante procedimientos óptimos en sentido de mínimo error cuadrático medio [8] y al campo del filtrado borroso, en el cual nos hemos centrado fundamentalmente en aspectos metodológicos de inferencia borrosa [9, 10], los cuales también han tenido aplicabilidad en la determinación de la maduración ósea en la infancia [15]. Asimismo, y fruto de la mencionada relación con el SPL, han sur-

gido interesantes desarrollos en modalidades emergentes de imagen médica [11, 12, 13].

3.3 Registado

Este frente, según consta en la memoria del proyecto, estaba inicialmente planteado para el subproyecto ULPGC. No obstante, fruto de la cooperación continuada de los grupos, y de la sinergia entre ambos que ello conlleva, en el subproyecto UVA se han hecho avances de interés también en esta línea. En particular, podemos mencionar el empleo inédito, hasta donde conocemos, de técnicas de registrado [14] para la resolución del problema de la maduración ósea en la infancia.

4 Conclusiones

Aparte de los objetivos previstos en el proyecto, los cuales entendemos cumplidos, la financiación conseguida nos ha permitido establecer relaciones con grupos nacionales e internacionales de enorme valor. En particular, el grupo investigador forma parte de una Red Temática de Investigación Cooperativa, financiada por el Fondo de Investigaciones Sanitarias, así como de una Red de Excelencia, financiada por la Comunidad Europea. Asimismo, de entre las tesis leídas durante este proyecto, una de ellas ha sido codirigida por el que redacta estas líneas y por un profesor del SPL, centro en el cual, adicionalmente, uno de nuestros investigadores ha realizado una estancia postdoctoral, financiada por la Comisión Fulbright. Finalmente, un investigador ha realizado una estancia en el Inria francés, financiada con una beca Marie-Curie.

Referencias

[1] C. Alberola, R. Cárdenes, M. Martín, M. A. Martín, M. A. Rodríguez, J. Ruiz, diSNei: A Collaborative Environment for Medical Images Analysis and Visualization, Lecture Notes in Computer Science, Vol. 1935, 2000, pp. 814-823.

[2] F. Simmross, N. Carranza, C. Palacios, Pablo Casaseca, M. A. Martín, S. Aja, J. Ruiz, C. F. Westin, C. Alberola, Group-Slicer: a Collaborative Extension of the 3D Slicer, Journal of Biomedical Informatics, (en prensa).

[3] M. Martín, C. Alberola, An Approach for Contour Detection of Human Kidneys from Ultrasound Images Using Markov Random Fields and Active Contours, Medical Image Analysis, Vol. 9. No. 1, 2005, pp. 1-23.

[4] C. Alberola, M. Martín, J. Ruiz. Comments on: A Methodology for Evaluation of Boundary Detection Algorithms on Medical Images, IEEE Trans. on Medical Imaging, Vol. 23, No. 5, Mayo 2004, pp.658-660.

[5] R. de Luis, M. Martín, J.I. Arribas, C. Alberola, A fully automatic algorithm for contour detection of bones in hand radiographs using active contours, Proc. of the IEEE Int. Conf. on Image Proc. ICIP-03, Barcelona, Septiembre 2003, Vol. III, pp. 41-424.

[6] R. San José, M. Martín, P.P. Caballero, C. Alberola, J. Ruiz, A theoretical framework to three-dimensional ultrasound reconstruction from irregularly-sampled data, Ultrasound in Medicine and Biology, Vol. 29, No. 2, Feb. 2003, pp. 255-269.

[7] R. San José, M. Martín, C. Alberola, J. Ellsemere, R. Kikinis, C. F. Westin, Freehand Ultrasound Reconstruction based on ROI Prior Modeling and Normalized Convolution, Lecture Notes in Computer Science, Vol. 2879, 2003, pp. 382-390.

[8] J. Ruiz, C. Alberola, C. F. Westin, Kriging Filters for Multidimensional Signal Processing, Signal Processing, 2005, Vol. 85, No. 2, pp. 413-439.

[9] S. Aja, C. Alberola, Fast Inference Using Transition Matrices, IEEE Trans on Fuzzy Systems, Vol. 22, No. 2. Abril 2004, pp. 170-182.

[10] S. Aja, C. Alberola, Fast Inference Using Transition Matrices: An Extension to Non-Linear Operators, IEEE Trans on Fuzzy Systems, (en prensa).

[11] M. Martín, C. Alberola, J. Ruiz, C. F. Westin, Regularization of Diffusion Tensor Maps using a Non-Gaussian Markov Random Field Approach, Lecture Notes in Computer Science, Vol. 2879, 2003, pp. 92-100.

[12] M. Martín, C. F. Westin, C. Alberola, 3D Bayesian Regularization of Diffusion Tensor MRI using Multivariate Gaussian Markov Random Fields, Lecture Notes in Computer Science, Vol. 3216, 2004, pp. 351-359.

[13] C.-F. Westin, M. Martín, C. Alberola, J. Ruiz and H. Knutsson, Tensor Field Regularization Using Normalized Convolution and Markov Random Fields in a Bayesian Framework, Visualization and Image Processing of Tensor Fields, Joachim Weickert and Hans Hagen (Eds), Springer, Berlin, 2005, (en prensa).

[14] M.A. Martín, E. Muñoz, M. Martín, C. Alberola, Articulated registration: elastic registration based on a wire-model, Proc. SPIE, Vol. 5747, 2005, pp. 182-191.

[15] S. Aja, R. de Luis, M. A. Martín, C. Alberola, A computational TW3 classifier for skeletal maturity assessment. A Computing with Words approach, Journal of Biomedical Informatics, Vol. 37, No. 2, Abril 2004, pp. 99 107.

MEDGENBASE: Acceso e integración virtual de bases de datos médicas y genéticas

López Alonso, V.; Vicente FJ.; Hermosilla, I.; Martín-Sánchez, F
Área de Bioinformática Médica. Instituto de Salud Carlos III
Ctra. Majadahonda a Pozuelo, Km. 2.
28220 – Majadahonda (Madrid)
Teléfono: 91 822 32 19 Fax: 91 509 79 17
E-mail: fmartin@isciii.es

The main goal of this project is the development of informatic tools that facilitate the integration of clinical and genetic data available in distributed, heterogeneous databases. It aims to unify this information in a virtual repository accesible through a single interface. A comprehensive analysis and selection of the most relevant genomic and disease-related information sources available in Internet has been carried out. The detailed knowledge on this subset of selected databases as well as several private databases have been shared with the other group participating in the project (UPM). The mapping of this heterogeneous databases has been done and the results of the unification process were validated. The final set of public and private databases, with genetic and medical content, including selected biomedical vocabularies and terminologies, constitute an integrated information space that can support information retrieval and data mining tasks in a biomedical research context.

1 Introducción

Uno de los grandes retos a los que se enfrenta la medicina en nuestros días consiste en integrar todo el nuevo conocimiento que se ha generado como resultado del Proyecto Genoma Humano, con los datos de la investigación en genómica y proteómica y los datos clínicos correspondientes a pacientes y enfermedades. Sólo de este modo será posible la obtención de nuevas soluciones diagnósticas, terapéuticas y preventivas con un impacto real en la salud de las personas. La integración de bases de datos heterogéneas implica diversos retos. En este proyecto se han considerado dos niveles de heterogeneidad: (1) plataformas tecnológicas diferentes (máquinas, sistemas operativos, y sistemas gestores de bases de datos), y (2) esquemas de base de datos diferentes. El punto (1) se refiere al aspecto de búsqueda de soluciones técnicas, mientras que el (2) se refiere a aspectos teóricos o conceptuales. El grupo del ISCIII ha desarrollado los aspectos englobados en (2) en relación al análisis de los conceptos de las bases de datos.

2 Métodos

Como primer paso, el sistema tiene como objetivo integrar diversas bases de datos privadas provenientes del entorno clínico (datos de pacientes, patología, datos de laboratorio, prescripción de medicamentos) con bases de datos de centros de investigación que recogen resultados de experimentos genómicos y proteómicos (secuencias, polimorfismos, expresión de genes obtenidos con microarrays, expresión de proteínas) proporcionando un modelo virtual de datos, con el que se puedan realizar por ejemplo, estudios de correlación entre las características genéticas y moleculares de una muestra con la evolución clínica de una enfermedad (Figura 1). Para centrar el ámbito de trabajo del sistema se decidió que los ámbitos gestionados por el sistema fueran inicialmente dos: cáncer y enfermedades genéticas raras. Se ha realizado un análisis cuidadoso de las bases de datos que integrará el sistema (Tabla 1). Se ha llevado a cabo una selección tanto de las bases de datos públicas disponibles en el campo de la genética y las enfermedades raras, asi como de las bases de datos privadas de las que disponiamos sobre enfermedades raras y cancer. La base de datos de tumores utilizada en la integración compilaba información sobre diversos aspectos de la historia clínica del paciente, tratamiento, anatomía patológica y análisis genético de las muestras de tumor.

En el entorno de las enfermedades raras, el sistema debería facilitar la recuperación de información relevante sobre las bases moleculares de una patología, accediendo a recursos públicos y facilitando los códigos y nombres de las entidades genéticas y clínicas subyacentes integrando terminologías estandarizadas, como:

- ICD–9–CM (International Classification of Disease, n, Clinical Modification)
- SNOMED (Systematized Nomenclature of Human and Veterinary Medicine)
- MeSH (Medical Subject Headings)
- UMLS (Unified Medical Language System).
- Gene Ontology (GO): Descripción de la función molecular, proceso biológico y componente celular
- UCL/HGNC: Human Gene Nomenclature: asigna un único símbolo a cada gen.
- MGED Ontology: Descripción de términos y datos de expresión génica.

Nuestro grupo ha colaborado intensamente en el proceso de “mapeo” que se utiliza para la creación de los esquemas virtuales individuales para cada base de datos. Este procedimiento requiere la intervención de un administrador que conozca bien y haya usado las fuentes y que pueda establecer las correspondencias entre el vocabulario compartido del dominio —

presente en el servidor de vocabulario— y las tablas y atributos que componen la base de datos. Del mismo modo, personal de nuestro grupo está validando el resultado del proceso de unificación que se ha usado para unir y mezclar varios esquemas virtuales, obteniendo un nuevo esquema virtual de unificación. Aunque dicha tarea se realiza de una manera totalmente automática, es necesario validar desde un punto de vista funcional los resultados que se obtienen.

3 Resultados

La existencia de las múltiples bases de datos genéticas y clínicas anteriormente descritas implica la existencia de distintas interfaces de usuario y de diversas metodologías de navegación. En este proyecto se han evaluado todas las vías de recuperación de los dos tipos de información y se han definido los requisitos funcionales del sistema para integrar los datos clínicos y genéticos. Esto permitirá al usuario navegar bajo una interfaz unificada, con una única metodología de búsqueda, basada en el empleo de ontologías o vocabularios controlados.

Cabe destacar, por último, que los dos grupos coordinados han participado conjuntamente en las tareas expuestas y participan en el proyecto europeo INFOGENMED y en la Red de Excelencia Europea en Informática Biomédica INFOBIOMED.

Agradecimientos

Esta investigación ha sido posible gracias al proyecto TIC2002-04444-C02-01 del Plan Nacional de Ciencia y Tecnología.

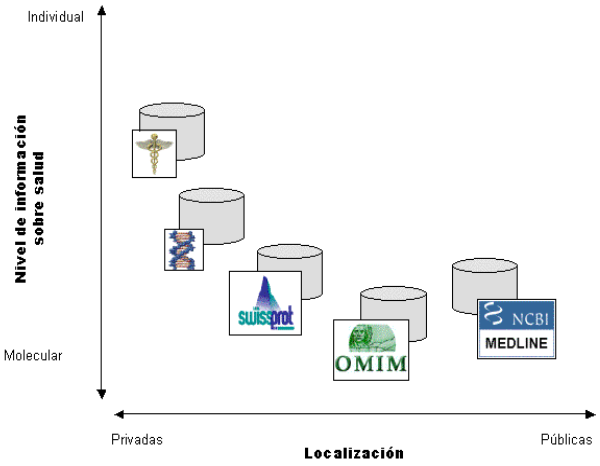


Figura 1. Espectro de bases de datos a considerar por el sistema

Referencias

Garcia Remesal M, Maojo V, Billhardt H, Crespo J, Alonso Calvo R, Perez D, Martin-Sanchez F, Sousa Pereira A, Babic A. ARMEDA II: Integrated Access to Heterogeneous Biomedical Databases. Proc. of the World Conference in Medical Informatics Medinfo 2004;1607. San Francisco, EE.UU.

Maojo V, Martin-Sanchez F. Bioinformatics: towards new directions for public health. Methods Inf Med. 2004;43(3):208-14.

M. García-Remesal, Victor Maojo, Holger Billhardt, José Crespo, R. Alonso-Calvo, D. Pérez-Rey, F. Martin-Sanchez, A. Sousa: ARMEDA II: Supporting Genomic Medicine through the Integration of Medical and Genetic Databases. Proc. of IEEE BIBE 2004: 227-236.

D. Pérez-Rey, V. Maojo, M. García-Remesal, R. Alonso-Calvo, H. Billhardt, F. Martin-Sánchez and A. Sousa. ONTOFUSION: Ontology-Based Integration of Genomic and Clinical Databases. Computers in Biology and Medicine. Aceptado para publicación.

Maojo, V., García-Remesal, M., Billhardt, H., Alonso-Calvo, R., Pérez, D., Calle, G. and Martin-Sanchez, F. Designing New Methodologies for Integrating Biomedical Information in Clinical Trials. Methods of Information in Medicine. En segunda revisión.

Alonso, R., Maojo, V., Billhardt, H., Martin-Sanchez, F., Garcia-Remesal, M. and Perez-del Rey, D. An Agent and Ontology-based System for Integrating Public Gene, Protein and Disease Databases. Journal of Biomedical Informatics. En segunda revisión.

Base de datos de genes y proteínas	Bases de datos sobre enfermedades raras	Centros diagnósticos, tests disponibles e investigación sobre medicamentos
EMBL - Secuencias de nucleótidos **dbSNP** - Variaciones puntuales del genoma	**CISATER** - Instituto de Investigación de Enfermedades Raras del Instituto de Salud Carlos III	**EDDNAL** - Directorio europeo de laboratorios de diagnóstico genético
OMIM - Enfermedades genéticas humanas **Entrez - Gene** - Información orientada a genes	**ORPHANET** - Base de datos del Instituto Nacional de Salud francés	**Genetests/GeneClinics** - base de datos de laboratorios estadounidenses e internacionales que realizan diagnóstico genético
KEGG - Catálogo de los genomas completos, procesos celulares, enzimas y reacciones enzimáticas	**NORD** - Base de datos privada de enfermedades raras	**PharmGKB** - investigación en farmacogenética
GeneCards - Enciclopedia de genes humanos	**FEDER** - federación española de enfermedades raras	**Clinicaltrials** – Ensayos clínicos
SwissProt - Secuencias de proteínas	**Eurordis** - Organización europea para Enfermedades Raras	
PROSITE - perfiles de proteínas		
InterPro - familias de proteína, dominios y sitios funcionales		

Tabla 1. Bases de datos seleccionadas para el proyecto

MEDGENBASE: Sistema de integración virtual de integración de información clínica y genómica a través de Internet

Maojo, V
Grupo de Informática Biomédica. Laboratorio de Inteligencia Artificial. Universidad Politécnica de Madrid

Resumen
El objetivo del proyecto ha sido la investigación de métodos y desarrollo de herramientas informáticas para la gestión e integración de información genética y médica, a través de Internet, utilizando un servidor de términos médico-genéticos y un sistema de integración basado en ontologías y agentes. Se ha realizado el desarrollo de una serie de herramientas informáticas para crear repositorios virtuales partiendo de conjuntos de bases de datos tanto públicas como privadas de carácter remoto y heterogéneo. El trabajo está basado en la idea de que un conjunto de bases de datos constituye un espacio de información, que puede ser descrito por uno o varios esquemas conceptuales construidos utilizando una terminología compartida y utilizado en tareas tales como búsqueda de información y análisis y minería de datos. Se han desarrollado varios métodos y herramientas que se presentan en el artículo.

Introducción
Al integrar bases de datos heterogéneas, se han considerado dos niveles de heterogeneidad: (1) plataformas tecnológicas diferentes (máquinas, sistemas operativos, y sistemas gestores de bases de datos), y (2) esquemas de base de datos diferentes. El punto (1) se refiere al aspecto de búsqueda de soluciones técnicas, mientras que el (2) se refiere a aspectos teóricos o conceptuales. Aunque bases de datos diferentes pueden contener información relativa al mismo dominio, usualmente tendrán diferentes esquemas. Por ello, es necesario crear un "mapping" o correspondencia entre el esquema físico de una base de datos y un esquema conceptual (en nuestro caso, basado en ontologías), creado mediante la utilización de terminología estándar asociada a un dominio (en nuestro caso, el dominio biomédico). La idea principal es que unidades de información similares almacenadas en diferentes bases de datos pueden ser nombradas de la misma manera en sus respectivos esquemas conceptuales de mapping. Este hecho permitiría la unificación automática de dos o más esquemas conceptuales que contengan entidades, creando un nuevo repositorio virtual de unificación que engloba la información de sus repositorios hijo. Los procesos de mapping y unificación permiten crear una organización taxonómica — es decir, una relación de contención o subsunción — del espacio de información disponible mediante la creación de una jerarquía de repositorios virtuales. Estos repositorios virtuales — ya sean de mapping o de unificación — facilitarían el acceso unificado a diferentes bases de datos remotas.

Métodos
Se está completando un módulo que se ha denominado "OntoFusión", que es un sistema de integración de bases de datos, utilizando vocabulario estándar para crear esquemas virtuales que representan cada fuente de datos que se desee integrar en el sistema. Dichos esquemas virtuales se emplean como interfaz con el usuario para navegar por las distintas bases de datos biomédicas y consultarlas de forma homogénea. Facilitando esta tarea al abstraer al usuario del sistema de gestión de base de datos empleado y de la estructura interna de la misma.

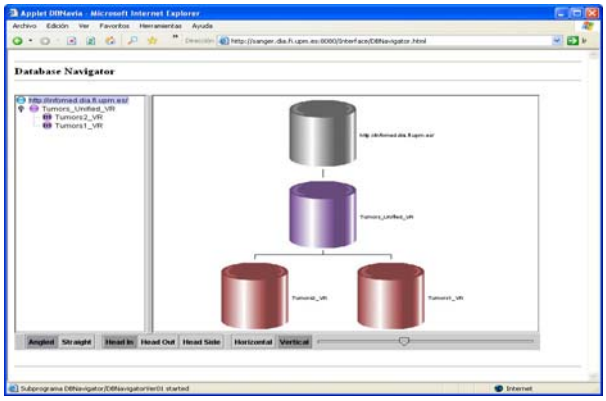
El Servidor de Vocabulario incluido en OntoFusión se utiliza como un subsistema de consulta para obtener términos estándar en el campo de la biomedicina. Estos términos se utilizan tanto a la hora de construir los esquemas virtuales (administrador) como al realizar consultas (usuario). Actualmente el sistema ONTOFUSION incluye 3 de las principales fuentes de términos en el campo de la biomedicina: *The Unified Medical Language System (UMLS), Gene Ontology* (GO) y *The Human Genome Nomenclature Commitee* (HGNC).

El proceso de mapping se utiliza para la creación de los esquemas virtuales individuales para cada base de datos. Es un procedimiento que requiere la intervención de un administrador. El administrador establece las correspondencias entre el vocabulario compartido del dominio —presente en el servidor de vocabulario— y las tablas y atributos que componen la base de datos. De esta manera se obtiene la lista de términos existentes en una base de datos —estos términos pueden corresponder a una parte de una tabla, a una tabla entera, o a un conjunto de tablas. Al finalizar el proceso de mapping se genera un fichero de correspondencias entre las tablas y atributos de la base de datos y términos del servidor de vocabulario, y otro fichero que describe el esquema virtual generado. Este segundo fichero generado se utiliza para generar el interfaz de usuario.

El proceso de unificación es usado para unir y mezclar varios esquemas virtuales, obteniendo un nuevo esquema virtual de unificación. Dicha tarea se realizará de una manera totalmente automática. Se ha desarrollado un algoritmo propio que mezcla dos o más esquemas virtuales —que contienen jerarquías de términos— y obtiene un esquema virtual unificado — con una nueva jerarquía contiendo todos los términos que aparecen en los esquemas de entrada. Los esquemas que se pueden unificar pueden ser tanto de las bases de datos reales como unificados, la única restricción que existe es que se hayan creado usando la misma jerarquía de términos.

Resultados preliminares

OntoFusión cuenta con un interfaz amigable, accesible por web para uso interno de pruebas, para realizar las consultas a las bases de datos integradas en el sistema. Inicialmente el usuario puede navegar entre los distintos esquemas virtuales, tanto unificados como no unificados, hasta determinar en cuál se encuentra la información que está buscando.



Navegador de Bases de Datos Virtuales

El sistema OntoFusión ha sido pensado para que sus distintas partes puedan ser ejecutadas de forma remota, de esta manera no sería necesaria su instalación en un servidor potente. Se ha migrado el sistema a la tecnología de agentes, gracias a la cual, la parte de comunicación a través de la red entre los distintos componentes se haría de forma transparente, gracias a las características de la programación orientada a agentes. Esta tecnología, ofrece al programador una comunicación fácil entre los distintos agentes de un sistema usando el lenguaje estándar de comunicación entre agentes ACL. Además la plataforma de agentes es la encargada de controlar el estado — activo o inactivo — y la situación de cada agente en cada momento.

La integración de datos requiere que se cubran las diferencias, tanto sintácticas como semánticas, que existen entre las distintas fuentes de datos biomédicas; y las ontologías son especialmente idóneas para tal tarea. Por ello en OntoFusión se está adaptando el sistema para sacar el mayor partido de esta nueva tecnología. Por un lado, los esquemas virtuales que representan las bases de datos que se deseen integrar se almacenarán utilizando ontologías. Por el otro, las fuentes del servidor de vocabulario se tratarán también como ontologías para facilitar su consulta y gestión. El lenguaje de representación de ontologías utilizado será DAML+OIL, basado en RDF y XML. Asimismo, OntoFusión será también compatible con el nuevo Ontology Web Language (OWL).

Cabe destacar, por último, que los dos grupos coordinados han participado conjuntamente en las tareas expuestas y participan en los proyectos europeos INFOGENMED y la Red de Excelencia Europea en Informática Biomédica INFOBIOMED, donde colaboran en diferentes tareas.

Publicaciones relacionadas con el proyecto

1. Garcia Remesal M, Maojo V, Billhardt H, Crespo J, Alonso Calvo R, Perez D, Martin-Sanchez F, Sousa Pereira A, Babic A. ARMEDA II: Integrated Access to Heterogeneous Biomedical Databases. Proceedings of the World Conference in Medical Informatics Medinfo 2004;2004:1607. San Francisco, EE.UU.
2. Maojo V, Martin-Sanchez F. Bioinformatics: towards new directions for public health. Methods Inf Med. 2004;43(3):208-14.
3. D. Pérez-Rey, Victor Maojo, M. García-Remesal, R. Alonso-Calvo: Biomedical Ontologies in Post-Genomic Information Systems. Proc of IEEE BIBE 2004: 207-214
4. M. García-Remesal, Victor Maojo, Holger Billhardt, José Crespo, R. Alonso-Calvo, D. Pérez-Rey, F. Martin, A. Sousa: ARMEDA II: Supporting Genomic Medicine through the Integration of Medical and Genetic Databases. Proc of IEEE BIBE 2004: 227-236
5. D. Pérez-Rey, V. Maojo, M. García-Remesal, R. Alonso-Calvo, H. Billhardt, F. Martin-Sánchez and A. Sousa. ONTOFUSION: Ontology-Based Integration of Genomic and Clinical Databases. Computers in Biology and Medicine. Aceptado para publicación
6. Maojo, V., García-Remesal, M., Billhardt, H., Alonso-Calvo, R., Pérez, D., Calle, G. and Martín-Sánchez, F. Designing New Methodologies for Integrating Biomedical Information in Clinical Trials. En segunda revisión en Methods of Information in Medicine.
7. Alonso, R., Maojo, V., Billhardt, H., Martin-Sanchez, F., Garcia-Remesal, M. and Perez-del Rey, D. An Agent and Ontology-based System for Integrating Public Gene, Protein and Disease Databases. Journal of Biomedical Informatics. En segunda revisión.

SIEMPRE: Seguimiento Inteligente y Extensible para el Modelado de la Práctica Educativa

Carlos Delgado Kloos, Abelardo Pardo Sánchez, José J. García Rueda,
Mª Carmen Fernández Panadero, Raquel M. Crespo, Sergio Gutiérrez, Luis de la Fuente
Departamento de Ingeniería Telemática. Universidad Carlos III de Madrid
Av. Universidad, 30. 28911 Leganés (Madrid). España
Teléfono: 91-624-8778. Fax: 91-624-8749. E-mail: cdk@it.uc3m.es

Abstract*. This paper summarizes the main achievements and tasks carried out within the SIEMPRE project. This project focuses on the tracking of students and the modelling of the educational process. The intelligent tools developed adapt the learning process according to the student's model, based on the data gathered by the system. The methodology and tools developed have been applied in real educational environments, proving their usefulness with experimental data. The objectives of the project have been accomplished successfully. The dissemination of results has been achieved by publishing at a good number of relevant international conferences and workshops. Several PhD theses have been started, and some of them actually finalized within the duration of the project.*

1 Introducción

El proceso educativo ha evolucionado en los últimos años, no sólo por la incorporación de las TIC, sino por su extensión a toda la vida profesional y la tendencia a centrarse en el alumno como protagonista. En este contexto, donde el seguimiento del alumno se considera clave para mejorar el rendimiento del proceso educativo, se hace imprescindible un modelo que formalice dicho seguimiento, extensible a lo largo de todo el proceso educativo, la adaptación de los contenidos en respuesta a este seguimiento y el desarrollo de herramientas que faciliten su aplicación a la práctica educativa.

2 Contribuciones técnicas

Una de las principales contribuciones del proyecto es la creación de EPM (*Educational Practice Model*), un modelo matemático para la caracterización y diagnóstico de procesos educativos. Basado en un riguroso análisis crítico de modelos precedentes, tecnológicos y pedagógicos, EPM constituye un acercamiento novedoso al problema del modelado educativo que permite:

- Caracterizar el curso en función de sus objetivos educativos
- Obtener medidas cuantitativas en las fases de diseño, impartición y evaluación de un curso, que permiten el diagnóstico y resolución de potenciales problemas
- Caracterizar el estado de conocimiento de un alumno, incluso en situaciones de incertidumbre.

La metodología desarrollada ha sido aplicada a diversos cursos, tanto presenciales como a distancia. Los datos obtenidos del seguimiento permiten, además, adaptar el proceso educativo al alumno, con vistas a la mejora del aprendizaje, mediante el uso de herramientas adecuadas. Dichas herramientas se han desarrollado en el proyecto en dos vertientes:

- Orientado a un escenario de trabajo individual, se ha desarrollado un sistema de tutoría inteligente vía web que selecciona los ejercicios que se plantean al alumno dependiendo de sus respuestas, basado en ejercicios parametrizables y grafos jerárquicos.
- Orientado a un escenario de trabajo colaborativo, se ha desarrollado un sistema de revisión entre iguales adaptativo, que personaliza el proceso de revisión en función del perfil del alumno, basado en clasificación borrosa y algoritmos genéticos.

Ambas herramientas se han utilizado en entornos educativos reales, concretamente en cursos de grado, como apoyo a la enseñanza presencial tradicional.

Otra de las contribuciones se centra en el proceso de generación de material educativo. Se ha definido e implementado una metodología para la creación flexible y reutilizable de contenidos que se puedan adaptar con facilidad a los requisitos de la práctica educativa. En términos concretos, se utiliza DocBook, una aplicación XML para la edición de documentación, como lenguaje en el que se generan los contenidos fuente, a partir de los cuales se produce de forma automática PDF para la impresión en papel o HTML para la presentación electrónica utilizando plantillas XSLT. Esto permite generar contenido independientemente de la presentación final, o, lo que es lo mismo, conseguir distintas presentaciones finales en función de parámetros dependientes del dispositivo de presentación o las características y avance del alumno y de los resultados de su seguimiento. La flexibilidad es muy grande, permitiéndose cambios en la fragmentación, la navegación, el aspecto, el orden de presentación o la selección de contenidos concretos a presentar.

3 Difusión de resultados

3.1 Publicaciones

Los artículos más importantes publicados en congresos internacionales son los indicados en las referencias [1] a [8]. Además se ha editado un libro en la editorial Kluwer correspondiente a un workshop [9].

3.2 Tesis doctorales

Durante la ejecución del proyecto, en concreto el 18 jun 2004, se ha leído una tesis doctoral, la de Mª Carmen Fernández Panadero (*EPM: Un modelo para la caracterización y diagnóstico de procesos educativos*). Otra tesis, la de Liliana Santacruz (*Automatización de los procesos para la generación, ensamblaje y reutilización de objetos de aprendizaje*) está a punto de leerse (ya se habrá leído cuando se publique este resumen). La tesis doctoral de Raquel Crespo se encuentra en la actualidad en un estado muy avanzado de realización, esperándose que se lea en este año. Las tesis doctorales de Sergio Gutiérrez y Pedro Muñoz Merino todavía están en desarrollo, aunque ya enfocadas.

3.3 Organización de eventos

En el marco del proyecto se ha organizado el Workshop *EduTech2004*, en Toulouse, 26-27 de Agosto de 2004 en el contexto del *World Computer Congress 2004*, patrocinado por "*IFIP WG 10.5 Design and Engineering of Electronic Systems*" en cooperación con "*IFIP WG 3.6 Distance Education*".

3.4 Relación con otros proyectos

Se participa como coordinador en el proyecto *E-LANE* del programa @lis de la UE. Los participantes en el proyecto son: UC3M, Telefónica I+D, GET-INT (Francia), U. Reading (UK), Trinity College Dublin (Irlanda), ITESM (México), U. Galileo (Guatemala), U. Cauca (Colombia), U Campinas (Brasil), U Chile (Chile). Este proyecto propone la integración de aplicaciones sobre la plataforma open source de *eLearning* .LRN y el diseño de una metodología de teleeducación, así como la integración de contenidos de las instituciones participantes. Este proyecto se centra en la demostración, con lo que la aportación del proyecto SIEMPRE a E-LANE se realiza desde el punto de vista de investigación y de innovación, donde los resultados del trabajo realizado en SIEMPRE se pueden poner en práctica. La sinergia entre ambos proyectos ha sido posible dada la proximidad de la temática, permitiendo aportar conocimiento y el resultado de las investigaciones realizadas.

4 Personal dedicado al proyecto

Los investigadores de plantilla que han trabajado en el proyecto son: Carlos Delgado Kloos (catedrático de universidad), Abelardo Pardo Sánchez (profesor titular de universidad), Mª Carmen Fernández Panadero (profesora titular de universidad interina), José J. García Rueda (ayudante doctor), Raquel Crespo García (ayudante). Sergio Gutiérrez Santos tiene una beca de FPI adjudicada en relación con el proyecto. Con cargo al proyecto se ha contratado sucesivamente a diversas personas: Mª Carmen Fernández Panadero (ahora profesora titular de universidad interina), Rosa Mª García Rioja (ahora becada en el CERN), Ralf Seepold (ahora profesor visitante), Arturo García Ares (ahora en la empresa privada), Enrique Sanchis Alcolea (ahora contratado con cargo a otro proyecto), Jonatan Tierno Alvite (ahora en la empresa privada) y Jesús Polo Torres. También han sido becados Pedro Muñoz Merino (ahora ayudante), Luis de la Fuente, David Vega Fontelos y José P. Escobedo del Cid.

Agradecimientos

Este trabajo ha sido realizado en el marco del proyecto TIC2002-03635 subvencionado por el Plan Nacional de I+D+i.

Referencias

[1] R.M. Crespo, A. Pardo, J.P. Somolinos Pérez, C. Delgado Kloos: *An Algorithm for Peer Review Matching Using Student Profiles (based on Fuzzy Classification and Genetic Algorithms)*, 18th Internat. Conf. on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems (iea/aie 2005). LNAI. Por publicar

[2] R.Mª Crespo, A. Pardo, C. Delgado Kloos: *An Adaptive Strategy for Peer Review*. 34th Frontiers in Education Conf., FIE 2004. Savannah, GA, USA. 20-23 oct. 2004

[3] C. Delgado Kloos, A. Pardo, P. Muñoz, N. Pérez: *A Type-Based Taxonomy of Items in Assessments*. En [9]

[4] S. Gutiérrez, A. Pardo, C. Delgado Kloos: *An Adapting Tutoring System Based on Hierarchical Graphs*, Adaptive Hypermedia AH, ago. 2004

[5] S. Gutiérrez, R.Mª García Rioja, A. Pardo, C. Delgado Kloos: *Beyond Simple Sequencing: Sequencing of Learning Activities using Hierarchical Graphs*, IASTED Internat. Conf. on Web-based Education WBE 04, Innsbruck, Austria, 16-18 feb. 2004

[6] R.Mª García Rioja, S. Gutiérrez, A. Pardo, C. Delgado Kloos: *A Parametric Exercise Based Tutoring System*, Frontiers in Education Conf., FIE 2003, Boulder, CO, EEUU. 5-8 nov. 2003

[7] A. Pardo: *A Multi-Agent Platform for Automatic Assignment Management*. Internat. Conf. on Innovation and Technology in Computer Science Education, jun. 2002

[8] A. Pardo: *A Platform for Parametrized Exercises in Web-Based Education*, Internat. Conf. Society for Information Technology and Teacher Education, mar. 2002

[9] C. Delgado Kloos, A. Pardo.: *EduTech: Computer-Aided Design meets Computer-Aided Learning*, Kluwer, Dordrecht 04, ISBN 1-4020-8161-8

TIC-204258-C03-01: Grid and peer-to-peer middleware for cooperative learning environments

Leandro Navarro-Moldes
Dep. d'Arquitectura de Computadors, Universitat Politècnica de Catalunya
Jordi Girona, 1-3, D6
08034 – Barcelona
E-mail: leandro@ac.upc.edu

***Abstract**. This paper presents the work done by the research group of the Universitat Politècnica de Catalunya within the coordinated project towards the definition of adequate grid and peer-to-peer middleware for CSCL (Computer Supported Collaborative Learning). Work has been centred on models based on autonomous decentralized and self-organized systems or "peer-to-peer" (P2P) systems for collaborative learning support. This support has been provided for synchronous and asynchronous collaboration, and based in platforms such as JXTA or Globus. Special emphasis has been put on combining information from different layers to optimize the organization of the system: based on the structure of the social network, based on the location of participants and the structure of the underlying network, based on the group structure. Finally, the evaluation has been done by simulation and limited prototypes. All these results were reflected in several publications in international journals, refereed conferences, as well as participation in new European projects.*

1 Introducción

Los sistemas distribuidos autónomos, descentralizados y auto-organizados basados en los modelos *peer-to-peer* [1] y Grid [2] ofrecen oportunidades para dar soporte informático a algunas formas de trabajo y aprendizaje cooperativo son potencialmente enormes. El diseño de estos sistemas precisa que tanto los participantes humanos como las máquinas se organicen de forma que se aprovechen las características, respeten las limitaciones, adapte y optimice el uso de las capacidades de humanos y máquinas trabajando juntos con estrecha interdependencia. Este proyecto ha estudia cómo estas nuevas formas de organizar los sistemas distribuidos pueden ser aprovechadas.

Para ello se ha investigado cómo se han de diseñar los elementos de un código intermedio o *middleware* que facilite que grupos de personas y las redes sociales que se forman, puedan colaborar. Se han identificado, propuesto y validado diversos componentes *middleware* que se benefician de las nuevas propuestas de Grid y *peer-to-peer*.

Este proyecto, que comenzó en Diciembre del 2002 y finaliza con Noviembre del 2005, se ha realizado en coordinación con equipos de investigación de la *Universidad de Valladolid* (grupo GSIC/EMIC de UVA), así como de la *Universidad Oberta de Catalunya* (UOC). Puede encontrarse información adicional en la Web en http://research.ac.upc.es/crac/

2 Objetivos concretos

El proyecto se ha enfocado hacia los siguientes objetivos:

- El análisis de las interacciones colaborativas en un espacio virtual de trabajo compartido.
- Definición de las características y componentes informáticos esenciales de una aplicación de soporte al aprendizaje cooperativo.
- Definición de una infraestructura "*middleware*" basada en los modelos de "Grid" y "P2P".
- Estudio y evaluación de mecanismos para la asignación de recursos y coordinación.
- Desarrollo y experimentación con aplicaciones concretas a partir del sistema propuesto.

3 Principales resultados

En el presente sub-proyecto se han conseguido los siguientes resultados principales:

- Se ha propuesto un algoritmo de consistencia rápida orientado por la demanda en sistemas distribuidos de gran escala que utilizan consistencia débil, como base para una aplicación de aprendizaje cooperativo a distancia. Este trabajo ha constituido la tesis de J. Jesús Acosta Elías, leída en Junio del 2003. [3]

- Se ha analizado y propuesto una infraestructura descentralizada para la colaboración. Este trabajo ha constituido la tesis doctoral de Joan Manuel Marquès i Puig, leída en Diciembre del 2003. [4]

- Se ha analizado y propuesto un marco conceptual y una propuesta de protocolo descentralizado para el desplegamiento de redes de aplicación en Internet. Este trabajo ha constituido la tesis de Oscar Ardaiz Villanueva, leída en Enero del 2004. [5]

- Se han estudiado las interacciones colaborativas en las redes sociales y se ha definido un protocolo *P2P* para redes de colaboración científica. Este trabajo ha constituido la tesis de José Mitre Silva, que se va a presentar durante el primer semestre del 2005. [6]

Estos resultados se han materializado en una gran variedad de resultados, incluyendo publicaciones, comunicaciones en congresos, así como ha dado la oportunidad de colaborar con Rededia S.L., una empresa *spin-off* del grupo de investigación de la UPC, organizar dos ediciones del taller CLAG (*International Workshop on Collaborative Learning Applications of Grid Technology*) en el marco de la conferencia CCGRID, participar en el proyecto Europeo Catnets (FET), entrar en el forum europeo de comunicación autonómica [7] y presentarse a algunos proyectos de la cuarta llamada de la UE, actualmente en fase de evaluación: la red de excelencia ACENET, el proyecto integrado SAICOM, y el proyecto STREP COSMA, sobre un entorno síncrono-asíncrono para aprendizaje cooperativo que la UPC lidera.

4 Conclusiones y trabajo futuro

En el sub-proyecto presentado en este artículo se ha planteado el problema de la creación de algunos elementos clave de un *middleware* adecuado para el apoyo a CSCL basado en tecnologías que siguen en la organización autónoma y descentralizada de las redes *peer-to-peer*, basado o inspirado en el paradigma de la computación grid. Entre sus principales aportaciones se destaca la definición de: mecanismos autónomos y descentralizados para el desplegamiento de redes de aplicación; para dar soporte a las tareas de colaboración de forma descentralizada basado en la distribución de eventos y replicación de objetos generados por grupos de personas; para distribuir cambios de forma eventualmente consistente en un entorno distribuido de gran escala; para formar redes superpuestas con estructura descentralizada que permiten a grupos de personas que colaboran compartir y encontrar objetos o documentos de interés para la comunidad.

Se ha visto que el diseño de sistemas que incluyen personas y máquinas que interactúan entre sí de forma intensa requiere poner atención y respetar aspectos sutiles pero esenciales tanto del funcionamiento del grupo humano como de la tecnología. La tecnología ofrece a la vez limitaciones y oportunidades que si se tienen en cuenta pueden aprovecharse para permitir nuevas oportunidades de colaboración entre personas que pueden estar dispersas respetando la autonomía del grupo.

Como trabajo futuro se plantea integrar los resultados del proyecto en un prototipo que permita realizar una evaluación con usuarios reales, así como incorporar los avances recientes en sistemas basados en modelos económicos y reputación, estudiar la posibilidad de extender el trabajo aplicándolo al nuevo concepto de computación autonómica [7] desarrollando los aspectos de auto-organización como reacción a la presencia de cambios ambientales, avanzar en el estudio de sistemas sensibles a la ubicación para detectar la emergencia de grupos en entornos de aprendizaje cooperativo síncrono.

Agradecimientos

El investigador principal de este proyecto quiere agradecer el trabajo realizado por todo el grupo de Sistemas Distribuidos de la Universitat Politècnica de Catalunya, a los grupos de la UVA y UOC con que nos hemos coordinado, así como por los becarios, colaboradores y alumnos que han participado en las experiencias educativas.

Referencias

[1] DS Milojicic, V Kalogeraki, R Lukose, K Nagaraja, “Peer-to-Peer Computing”, HP Laboratories Palo Alto, Technical Report HPL-2002-57.

[2] I. Foster, C. Kesselman, S. Tuecke. "The Open Grid Services Architecture". In: *The Grid 2: blueprint for a future computing infrastructure*, eds. I. Foster, C. Kesselman. San Francisco, CA, USA: MKn Publishers (2004).

[3] J. Jesús Acosta Elías. “Algoritmo de consistencia rápida orientado por la demanda en sistemas distribuidos de gran escala”, Tesis doctoral, 2003.

[4] Marquès, J.M. “LaCOLLA: una infrastructura autònoma i autoorganitzada per facilitar la col·laboració”. Tesis doctoral, 2003.

[5] Oscar Ardaiz Villanueva, “Protocolo descentralizado para el desplegamiento de redes de aplicación en Internet”. Tesis doctoral, 2004.

[6] José Mitre Silva, “Protocolo *peer-to-peer* para redes de colaboración científica”. Tesis doctoral, 2005.

[7] Autonomic Communication Forum; url: http://www.autonomic-communication.org/

TIC-204258-C03-02: Grid and peer-to-peer middleware for cooperative learning environments

Yannis A. Dimitriadis
Dpto. de Teoría de la Señal, Comunicaciones e Ingeniería Telemática. Universidad de Valladolid
ETSI de Telecomunicación. Camino Viejo del Cementerio s/n
47011 – Valladolid (Valladolid)
E-mail: yannis,@tel.uva.es

Abstract*. This paper presents the work done by the research group of the University of Valladolid within the coordinated project towards the definition of adequate grid and peer-to-peer middleware for CSCL (Computer Supported Collaborative Learning). The main proposal consists in a new system called Gridcole, based on the specific technology of grid services that offers support for tailorability, design and interpretation of collaboration scripts, as well as access to supercomputing and special hardware facilities. Through analysis of the domain, we advanced in the detection and formalization of essential components and structures that appear at educational best practices, called Collaborative Learning Flow Patterns using the IMS-LD standard. Finally, the design, set-up, realization and evaluation of prototypes within significative collaborative learning scenarios provided a validation of the proposed solutions. All these results were reflected in several publications in international journals, refereed conferences, as well as participations to new european projects.*

1 Introducción

El aprendizaje colaborativo apoyado por ordenador o CSCL constituye un nuevo paradigma referente al uso de las TIC en educación que ha atraído mucha atención por parte de tecnólogos y educadores. Entre sus principales características podemos mencionar el uso de las interacciones sociales como medio para el aprendizaje, el diseño participativo entre todos los actores involucrados, así como el carácter distribuido de las herramientas y sistemas que lo apoyan [1].

Por otro lado, se ha detectado la necesidad de forma clara de construir unas capas de *middleware* que apoyen las aplicaciones CSCL y que permitan reutilización, flexibilidad, adaptación etc. por parte de educadores y tecnólogos. Esta capa aumentaría la eficiencia en el proceso de ingeniería de aplicaciones distribuidas. La aparición de nuevas propuestas de organización de sistemas distribuidas, tales como computación en malla (*grid*) [2] o entre pares (*peer-to-peer*, P2P), ha suscitado la necesidad de estudiar su viabilidad en sistemas distribuidos de aprendizaje y de CSCL en particular.

Por todo ello, en este proyecto se han planteado los objetivos de proponer un sistema de middleware basado en los dos paradigmas mencionados anteriormente, de analizar los elementos esenciales de las aplicaciones CSCL, de poner en marcha dichas aplicaciones significativas y ofrecer sistemas de análisis de interacciones y de evaluación de su uso en entornos reales. Este proyecto se ha realizado en coordinación por equipos de investigación de la *Universidad de Valladolid* (grupo GSIC/EMIC de UVA en el presente subproyecto), así como de la *Universitat Politècnica de Catalunya* (UPC) y de la *Universitat Oberta de Catalunya* (UOC).

2 Principales resultados

En el presente subproyecto se han conseguido los siguientes resultados, que se reflejan de forma parcial en el esquema de uso de la Figura 1.

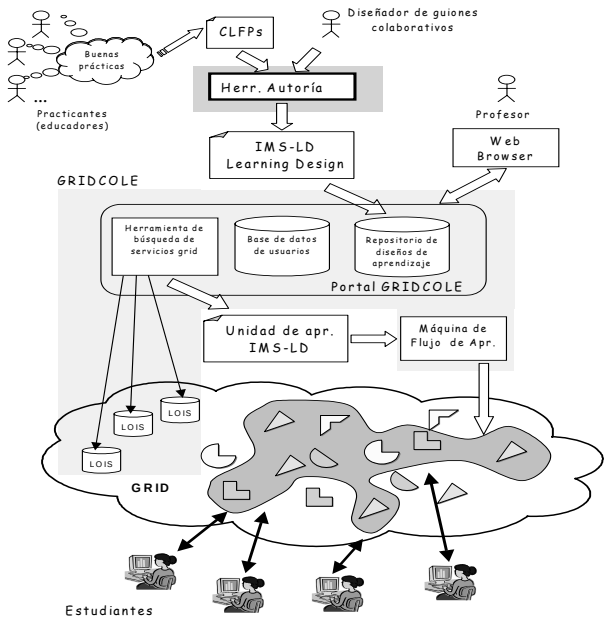


Figura 1: Esquema global de la solución propuesta

- Se ha analizado y validado la adecuación de la tecnología de servicios grid que se apoyan en el paradigma grid para situaciones de CSCL [3].

- Se ha propuesto el sistema Gridcole como núcleo de puesta en marcha de aplicaciones CSCL [4].

Este sistema tiene la propiedad de maleabilidad (*tailorability*), y da la posibilidad de uso de nuevos escenarios de uso, ofrece acceso a recursos supercomputacionales o de hardware específico, al mismo tiempo que permite interpretar y ejecutar guiones de aprendizaje.

- Por otro lado, se propone un sistema de creación de diseños de aprendizaje apoyado por el estándar de IMS-LD, donde se permite al educador estructurar las actividades a realizar por los participantes junto con los recursos asociados en guiones (*scripts*) de aprendizaje. Estos guiones se almacenan en el repositorio de diseños de aprendizaje de Gridcole, se interpretan por la máquina de flujo de aprendizaje y se asocian a herramientas en forma de servicios grid que se escogen entre los ofrecidos por los proveedores.

- Para facilitar el diseño participativo, la creación de escenarios significativos de uso y promocionar la flexibilidad y la reutilización de software, se ha llegado a la propuesta de los Patrones de Flujo de Aprendizaje Colaborativo (CLFPs) [5]. Estos patrones formalizan buenas prácticas de organización de aprendizaje pudiendo definir de esta forma flujos adecuados y definir los elementos que están presentes en aplicaciones típicas de CSCL.

- Por último, y a través de la realización de prototipos de software y de su uso en escenarios reales, se ha avanzado en la propuesta de un marco para el análisis de las interacciones producidas en las actividades colaborativas, la coordinación a nivel de objeto, y la influencia de los roles en las tareas de evaluación y apoyo integrado [6].

Estos resultados se han materializado en un capítulo de libro, 1 premio al mejor artículo de la conferencia ICALT 2004, 2 artículos invitados en revistas internacionales, 2 artículos publicados en revistas incluidas en el *ISI-JCR*, 3 artículos publicados en otras revistas internacionales, 4 artículos publicados en la serie *LNCS* de *Springer Verlag*, 24 comunicaciones en congresos, la co-organización de 2 *workshops* dentro de la conferencia IEEE CCGrid (GLAG 2004 y CLAG+.Edu 2005), así como la participación en la red de excelencia *Kaleidoscope* del programa de IST y un proyecto del programa *e-learning* de la Unión Europea.

3 Conclusiones y trabajo futuro

En el subproyecto presentado en este artículo se ha planteado el problema de la creación de un *middleware* adecuado para el apoyo a CSCL basado en tecnologías que siguen en paradigma de la computación grid. Entre sus principales aportaciones se destaca el sistema maleable Gridcole que permite la interpretación de diseños de aprendizaje, la búsqueda de herramientas asociadas y su integración usando servicios grid. Por otro lado, se ha propuesto el uso de patrones de flujo de aprendizaje colaborativo (CLFPs) que recogen buenas prácticas y las formalizan con el estándar IMS-LD, contribuyendo en el mejor uso del diseño participativo, la flexibilidad y la reutilización del software orientado a servicios. Finalmente, se ha contribuido a la mejora del análisis de las interacciones y su uso para el apoyo al aprendizaje y la evaluación.

En los siguientes pasos se plantea, entre otros, la evaluación de Gridcole, la propuesta de una ontología de servicios para una mejor búsqueda de los mismos, así como la integración de los mecanismos de apoyo y de evaluación usando el concepto de rol.

Agradecimientos

El investigador principal de este proyecto quiere agradecer el trabajo realizado por todo el grupo GSIC / EMIC de la Universidad de Valladolid, así como por los becarios, colaboradores y los alumnos que han participado en las experiencias educativas.

Referencias

[1] P. Dillenbourg. Collaborative Learning: cognitive and computational approaches, Oxford, UK: Elsevier Science (1999).

[2] I. Foster, C. Kesselman, S. Tuecke. "The Open Grid Services Architecture". In: *The Grid 2: blueprint for a future computing infrastructure*, eds. I. Foster, C. Kesselman. San Francisco, CA, USA: Morgan Kaufmann Pub. (2004).

[3] M.L. Bote, Y.A. Dimitriadis, E. Gómez, "Grid Characteristics and uses: a grid definition". Lecture Notes in Computer Science, Springer-Verlag 2970 (2004) 291-298.

[4] M. Bote, L.M. Vaquero, G. Vega, Y. Dimitriadis, J.I. Asensio, E. Gómez, D. Hernández, "A tailorable collaborative learning system that combines OGSA grid services and IMS-LD scripting". Lecture Notes in Computer Science, Springer Verlag 3198 (2004) 305-321.

[5] D. Hernández, J.I. Asensio, Y.A. Dimitriadis, "IMS Learning Design support for the formalization of collaborative learning patterns", IEEE Educational Technology and Society, 8 (2005) (en prensa).

[6] A. Martínez, Y.Dimitriadis, E. Gómez, B. Rubia, P. de la Fuente, "Combining qualitative and social network analysis for the study of classroom social interactions", Computers and Education, , 41 (2003) 353-368.

TIC-204258-C03-03: Grid and peer-to-peer middleware for cooperative learning environments

Atanasi Daradoumis
Estudios de Informática y Multimedia. Universitat Oberta de Catalunya
Av. de Tibidabo 39-43
08035 – Barcelona
E-mail: adaradoumis@uoc.edu

***Abstract**. This paper presents the work done by the research group of the Open University of Catalonia within the coordinated project towards the definition of adequate grid and peer-to-peer middleware for cooperative learning environments. The main results of our research consist in designing a layered framework for modelling online collaborative learning interactions which in turns leads to the development of a computational platform, called Collaborative Learning Purpose Library (CLPL), that can be used as a basis for the development of Collaborative Learning applications that achieve an efficient event (group activity) information management. To that end, we developed a Grid approach in order to increase the efficiency of processing event log files. Finally, we developed a middleware architecture, called LaCOLLA, that follows the peer-to-peer paradigm and can be used to build collaborative applications that provide general purpose collaborative functionalities while it pays special attention to the autonomy of its members and to self-organization of the components of the infrastructure. All these results were reflected in several publications in international journals, refereed conferences, as well as participations to new European projects.*

1 Introducción

El aprendizaje colaborativo apoyado por ordenador (CSCL) ha hecho surgir varias cuestiones que necesitan ser respondidas de manera eficaz. Estas cuestiones están relacionadas, por una parte, con la investigación sobre el análisis de la interacción y, por otra parte, con el desarrollo de sistemas grid y middlewares colaborativos. Estas aproximaciones se han orientado a identificar y a explorar los factores que afectan la eficacia y el éxito del trabajo y del aprendizaje de grupos online [1].

Inicialmente se detectó la necesidad de diseñar aproximaciones para el análisis de las complejas interacciones que suceden en un grupo de aprendizaje colaborativo online con el objetivo de monitorizar, evaluar y dar apoyo a los diferentes participantes (alumnos y profesores) [1]. Además, la aparición de nuevas propuestas de organización de sistemas distribuidas, tales como computación en malla (*grid*) [2] o la definición de arquitecturas en capas de middleware tipo peer-to-peer que apoyan las aplicaciones CSCL han suscitado la necesidad de estudiar su viabilidad en CSCL.

Por todo ello, la contribución de la UOC en el proyecto se ha planteado en dos partes. En una de las partes se ha planteado como objetivo proponer un marco para la modelización de interacciones de aprendizaje colaborativas online, el cual ha dado lugar a la construcción de una plataforma para el desarrollo de aplicaciones colaborativas y al uso del Grid para aumentar el procesamiento de la información que resulta de estas aplicaciones. Por otro lado, se ha propuesto un middleware peer-to-peer para facilitar la construcción de aplicaciones colaborativas para grupos dispersos en Internet que funcionen de manera autónoma y autoorganizada. Este proyecto se ha realizado en coordinación con equipos de investigación de la *Universidad Oberta de Catalunya* (UOC) (en el presente subproyecto), así como de la *Universidad de Valladolid (UVA)* y de la *Universitat Politècnica de Catalunya* (UPC).

2 Principales resultados

En el presente subproyecto se han conseguido los siguientes resultados:

- Se ha propuesto un marco metodológico para el aprendizaje colaborativo basado en proyectos en entornos virtuales [3].

- Se ha diseñado un modelo de evaluación del aprendizaje colaborativo de cursos online de la UOC en varios niveles y aspectos [4].

- Se ha desarrollado un modelo conceptual de apoyo al análisis de las interacciones colaborativas, definiendo y clasificando en categorías importantes y variados indicadores de la colaboración [5].

- Se han definido varios métodos concretos de análisis de las interacciones de naturaleza complementaria con la colaboración de la Universidad de Valladolid (UVA) con el propósito de evaluar las experiencias colaborativas efectuadas en la UOC. Además se

realizó un estudio comparativo de varios métodos de análisis sobre los mismos datos de interacción y se avanzó hacia un sistema integrado de análisis de interacciones [5].

- Se ha propuesto una plataforma genérica CSCL basada en componentes de software para la construcción de aplicaciones CSCL específicas que proporcionan una gestión eficiente de la información generada por las interacciones colaborativas [6].

- Se ha estudiado los requisitos de las aplicaciones colaborativas a escala Internet que dan soporte a grupos colaborativos autoorganizados y descentralizados utilizando los recursos aportados por los participantes en la colaboración [9]. A partir de estos requisitos, se ha diseñado e implementado un prototipo de middleware peer-to-peer, LaCOLLA, que recoge los principales aspectos de la arquitectura propuesta [7,9].

- Por último, se ha analizado y validado la adecuación de la plataforma genérica CSCL así como el aumento de la eficiencia del procesamiento de ficheros log de la actividad grupal que se apoya en el paradigma Grid [6, 8].

Estos resultados se han materializado en 1 tesis doctoral, en 1 capítulo de libro, 3 artículos publicados en la serie *LNCS* de *Springer Verlag*, 8 comunicaciones en congresos, así como la participación en la red de excelencia *Kaleidoscope* del programa de IST y un proyecto, Virtual Math Teams Project (VMT), de Drexel University, Philadelphia, USA.

3 Conclusiones y trabajo futuro

En el subproyecto presentado en este artículo se planteó la creación de un marco para la modelización de interacciones de aprendizaje colaborativas online, la construcción de una plataforma para el desarrollo de aplicaciones colaborativas y en el uso del Grid para aumentar el procesamiento de la información que resulta de estas aplicaciones. Además, se propuso un sistema middleware peer-to-peer autónomo y autoorganizado para facilitar la construcción de aplicaciones colaborativas que den apoyo a grupos dispersos en Internet.

En los siguientes pasos se plantea, entre otros, el desarrollo de: un modelo multi-dimensional de análisis de interacciones colaborativas que incluya diferentes enfoques, de herramientas colaborativas, y de un prototipo basado en la tecnología Grid para procesar y analizar la información de la actividad de los grupos de trabajo. Además se plantea la propuesta de algoritmos peer-to-peer de planificación, optimización y asignación de recursos y servicios en entornos de aprendizaje basados en sistemas económicos, de comportamiento optimista y de localización estructurada y que son válidos para sistemas descentralizados y autónomos.

Agradecimientos

El investigador principal de este subproyecto quiere agradecer el trabajo realizado por todo el grupo Distributed Systems & CSCL de la Universitat Oberta de Catalunya, así como por los becarios, colaboradores y los alumnos que han participado en las experiencias educativas.

Referencias

[1] P. Dillenbourg. Collaborative Learning: cognitive and computational approaches, Oxford, UK: Elsevier Science (1999).

[2] I. Foster, C. Kesselman, S. Tuecke. "The Open Grid Services Architecture". In: *The Grid 2: blueprint for a future computing infrastructure*, eds. I. Foster, C. Kesselman. San Francisco, CA, USA: MKn Publishers (2004).

[3] Daradoumis T., and Xhafa, F. Problems and Opportunities of Learning together in a Virtual Learning Environment. In: Computer-Supported Collaborative Learning in Higher Education, Chapter 11. Roberts, T. S. (Eds). Idea Group Press, pp. 218-233 (2004).

[4] Daradoumis T., Xhafa, F. and Marquès J.M. Evaluating Collaborative Learning in a Virtual Groupware Environment. Proc. of the IASTED Int. Conference on Computers and Advanced Technology in Education, pp. 438-443 (2003).

[5] Daradoumis T., Martínez A., and Xhafa, F. An Integrated Approach for Analysing and Assessing the Performance of Virtual Learning Groups. Lecture Notes in Computer Science (3198). Springer-Verlag, pp. 289-304 (2004).

[6] Caballé, S., Xhafa, F., Daradoumis, T. and Marquès J.M. Towards a Generic Platform for Developing CSCL Applications Using Grid Infrastructure. In: Proc. of the 1st Int. Workshop on Collaborative Learning Applications of Grid Technology (CLAG), IEEE/ACM/IEEE Computer Society, (2004).

[7] Marquès J.M., Navarro L. and Daradoumis, T. Extending the Scope of Asynchronous Collaboration: a Matter of Being Autonomous and Self-sufficient. In: Proc. of the 13th IEEE Int. Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises, IEEE Computer Society, pp.87-92 (2004).

[8] Xhafa, F., Caballé, S., Daradoumis, T. & Zhou, N. A Grid-Based Approach for Processing Group Activity Log Files. Lecture Notes in Computer Science (LNCS 3292) Springer-Verlag, pp. 175-186 (2004).

[9] Marquès, J.M. LaCOLLA: una infrastructura autònoma i autoorganitzada per facilitar la col·laboració. Tesis doctoral. 2003.

eCLUB: Evolución de un entorno de enseñanza basado en escritorio hacia la computación ubicua. Aplicación a la enseñanza de materias experimentales

A. I. Molina, M. A. Redondo, C. Bravo, M. Ortega
Grupo de Investigación CHICO. Departamento de Informática. Universidad de Castilla-La Mancha
Escuela Superior de Informática. Paseo de la Universidad, 4.
13071 - Ciudad Real
E-mail: { AnaIsabel.Molina, Miguel.Redondo, Crescencio.Bravo, Manuel.Ortega}@uclm.es

Abstract*. The main goal of this project is to incorporate the ubiquitous computing paradigm in the teaching and learning of domains with a high experimental degree in order to take into account mobile computing possibilities. We will take as starting point a collaborative e-learning environment based on the desktop metaphor, called "DomoSim-TPC". To achieve our goal, we have analysed the tasks (already modelled in the aforementioned system), which are susceptible of improvement through ubiquitous computing. Once these tasks defined, we have developed a flexible architecture that will support them and will be extensible and applicable to other situations and necessities. With this architecture we have implemented a prototype materialising the theories outlined. This prototype will be applied to the learning of Domotics and integrated in DomoSim-TPC environment.*

1 Introducción

La aparición de una nueva generación de dispositivos móviles, cada vez más portables, potentes y asequibles, junto con los avances que se están produciendo en el ámbito de las telecomunicaciones, está llevando al uso de este tipo de dispositivos en ámbitos para los cuales no se pensó inicialmente. Uno de los campos que se pueden ver más beneficiados de este hecho es el de la educación, originando cambios no solo en el modo en que la información es usada y compartida, sino en la propia forma de operar. Pero la plena introducción de estos dispositivos no se apoya tanto en el desarrollo y abaratamiento de los mismos, sino en la creación de un software que dé soporte a nuevas aplicaciones.

El presente artículo se centra en exponer el proceso de evolución del entorno Domosim-TPC, un entorno telemático CSCL para la enseñanza de la domótica basado en escritorio hacia la computación móvil. El artículo se estructura de la siguiente forma: en primer lugar se hace una introducción al concepto de computación ubicua, las principales características de la herramienta de partida (Domosim-TPC) y algunos conceptos sobre el modelado de tareas. A continuación se describen los pasos seguidos para hacer evolucionar Domosim-TPC hacia la computación móvil, y los resultados obtenidos. Finalmente, se apuntan algunas conclusiones que se obtienen de este trabajo y futuras líneas de actuación.

2 Antecedentes

La Computación Ubicua como paradigma de interacción [1, 2] desplaza el concepto de uso de la computadora distribuyendo múltiples computadoras de poca potencia a lo largo del entorno, y tratando de ocultar su presencia y utilización. Se apuesta por la implantación de este nuevo paradigma en el aula con fines educativos [3], presentando algunos de los beneficios pedagógicos que aportará su utilización.

El sistema DomoSim-TPC es un entorno telemático para el aprendizaje colaborativo del diseño de instalaciones domóticas. Esta herramienta da soporte a la realización de actividades de resolución de problemas en grupo mediante la planificación, diseño y la simulación de forma distribuida de las soluciones planteadas. Este sistema se basa en los paradigmas de interacción asistida abstracta de conceptos [4] y de manipulación directa; y está basado en la metáfora de escritorio. Fue sometido a un proceso de evaluación formativa, cuyas conclusiones sirven de motivación para este trabajo [5 , 6].

3 Evolución de Domosim-TPC hacia la computación ubicua

El **objetivo** perseguido es la incorporación del paradigma de computación móvil (encuadrado dentro de la computación ubicua), a un entorno colaborativo de e-Learning (Domosim-TPC). En particular, algunas tareas a las que da soporte esta herramienta pueden resultar más reales, accesibles y motivantes si se abordan incluyendo características propias de la computación móvil.

El proceso de evolución de Domosim-TPC hacia la computación ubicua ha pasado por varias **etapas**: (1) *Análisis de las tareas* que sean susceptibles de ser *mejoradas* mediante computación ubicua. (2) *Replanteamiento del modelado y diseño* de las tareas seleccionadas para adaptarlas al nuevo paradigma. (3)

Implementación de un prototipo que materialice las teorías que planteamos. (4) *Evaluación el prototipo* en contextos reales. (5) Identificación de *patrones de tareas* que puedan ser comunes *en entornos colaborativos de enseñanza* basadas en la resolución de problemas.

Uno de los principales problemas del proceso de evolución planteado es la *adaptación de la interfaz de usuario* al nuevo dispositivo, teniendo en cuenta las limitaciones que este plantea. Se propone un método que permita automatizar dicho proceso, el cual sería aplicable a cualquier entorno CSCL, sea cual sea el dominio de aplicación en el que nos encontremos. Para ello se hace uso del *modelo de tareas* de la aplicación. Existen varios métodos para el análisis de tarea, que se diferencian en el grado de formalismo y finalidad. Hemos elegido la técnica CTT (*ConcurTaskTrees)* [7] para el modelado y remodelado de nuestro sistema, por ser la que mejor se adapta a nuestras necesidades (notación gráfica, fácil de entender, estructura jerárquica que permite representar diferentes niveles de abstracción y refinamientos progresivos, amplio juego de operadores temporales, etc.). Hemos demostrado que aunque CTT permite el modelado de aplicaciones cooperativas no permite el correcto modelado de aplicaciones colaborativas; aspecto que supone una nueva línea de investigación resultante de este trabajo.

El resultado del proceso de adaptación ha dado lugar a la creación del prototipo llamado Domosim-Mob [8], que se encuentra actualmente en proceso de evaluación.

Una de las ventajas que se extraen del modelado genérico de las tareas es la posibilidad de *generación automática de interfaces* adaptados a cada uno de los dispositivos de los que se puede hacer uso en nuestra propuesta de clase ubicua (PDA, PC, Smartphone,...). Partiendo de la descripción abstracta de estas tareas genéricas se pueden obtener interfaces concretas, y cuya realización esté justificada en el dispositivo a emplear, dentro del modelo de la clase. El análisis de tareas del que se parte debe hacerse a tan bajo nivel que quede especificado el *tipo de interacción* que se realiza en cada caso (entrada de texto, selección de un valor booleano, elección de un valor numérico, finito o infinito). Contar con este tipo de información facilitará la *elección del widget* concreto que mejor de soporte a dicha tarea.

4 Conclusiones

En este artículo hemos expuesto nuestro objetivo de incorporar el paradigma de computación ubicua a un entorno colaborativo de e-Learning basado en la metáfora de escritorio de cara a explotar las posibilidades de la informática móvil en la enseñanza de dominios con un alto grado experimental. Al evaluar el sistema pretendemos ir más allá, intentando la *generalización de nuestros planteamientos* a otros dominios y situaciones. Dicha generalización pasa por la identificación de ciertos *patrones de interacción de naturaleza colaborativa y ubicua* y la definición de una serie de *líneas guía*, que sirvan de base a la *generación automática de interfaces de usuario CSCL* para los distintos dispositivos móviles presentes en nuestro modelo de clase ubicua [3]. Dicho proceso se hace partiendo del diseño conceptual de la correspondiente aplicación basada en escritorio, y es independiente del dominio de aplicación concreto.

Agradecimientos

Este trabajo ha sido realizado gracias al apoyo prestado por la Junta de Comunidades de Castilla-La Mancha y por Ministerio de Ciencia y Tecnología en el marco de los proyectos PBI-02-026 y TIC2002-01387 respectivamente.

Referencias

[1] Weiser, M., The computer for the twenty-first century. *Scientific American*, 1991: p. 94-104.

[2] Weiser, M., The future of Ubiquitous Computing on Campus. *Comm. ACM*, 1998. 41-1.

[3] Ortega, M., et al., Ubiquitous Computing and Collaboration: New Paradigms in the classroom of the 21st Century., *Computers and Education: Towards a Interconnected Society*, M.O.a.J. Bravo, Editor. 2001, Kluwer Academic Publishers. p. 261-273.

[4] Bravo, J., et al. Interacción asistida abstracta en la clase ubicua. *Revista Iberoamericana de Inteligencia Artificial*. 2001. núm. 16, pp. 49-54.

[5] Redondo, M.A., Planificación Colaborativa del Diseño en Entornos de Simulación para el Aprendizaje a Distancia, Phd Thesis in Departamento de Informática. 2002, Universidad de Castilla-La Mancha: Ciudad Real. p. 334.

[6] Bravo, C., Un sistema de Soporte al Aprendizaje Colaborativo del Diseño Domótico Mediante Herramientas de Modelado y Simulación., Phd Thesis in Dpto. de Informática. 2002, Universidad de Castilla - La Mancha.

[7] Paternò, F., C. Mancini, and Meniconi. ConcurTaskTree: A diagrammatic notation for specifying task models. in IFIP TC 13 International Conference on Human-Computer Interaction Interact'97. 1997. Sydney: Kluwer Academic Publishers.

[8] Molina, A.I., Redondo, M.A., Ortega, M., Bravo, C., (In Press). A system to support asynchronous collaborative learning tasks using PDAs. *Journal of Universal Computer Science*. Know-Center.

Sistema Seguro para la Certificación Remota de Documentos (CREDO: TIC2002-00249)

Rico-Novella, F.J., Forga Alberich, J, Sanvicente Gargallo, E. Cruz Llopis, L.J. Alins Delgado, J., Mata Díaz, J.
Departamento de Ingeniería Telemática. Universidad Politécnica de Cataluña
E-mail: f.rico@entel.upc.es

***Abstract**. The patented CredoTicket solution allows the generation of consumable digital documents (tickets and electronic coupons) that can be validated off-line, avoiding reutilization in a secure and automatic setting. The ticket generation is achieved by CredoTicket servers using highly robust cryptographic techniques that prevent forging. Clients receive and handle tickets as printable PDF documents (that incorporate barcodes) or as simple numeric data. Each generated ticket is associated to a unique smart card belonging to the client. The ticket checking and validation is carried out by CredoTicket readers. The readers read the ticket code and check their validity interacting with the smart card. To perform this task, readers do not need to communicate with either the server that generated the ticket or with any other system or reader (off-line checking). Once validated, the ticket is recorded in a list kept in the smart card to prevent further utilization*

1 Introducción

En la actualidad nadie duda de las ventajas de la venta de tickets online y la impresión de los mismos en casa del cliente. Para el comprador, este tipo de venta resulta más cómoda, ya que evita colas o tiempos de espera excesivos que pueden desalentar la posible compra. Para el vendedor, aumenta de forma sustancial el número de posibles compradores, al mismo tiempo que reduce los costes, puesto que no es necesario disponer de complejos y costosos dispositivos dispensadores o, en su defecto, personal en taquillas.

Sin embargo, este nuevo tipo de venta puede acarrear problemas de seguridad. El alto valor de algunos tickets hace que la falsificación de los mismos sea cada vez más atractiva, a la vez que los mecanismos de falsificación son más baratos y sofisticados. De hecho, los sistemas usuales de impresión en casa del cliente permiten fácilmente la duplicación de los tickets impresos. Ello obliga a que todos los controles de entradas de un evento estén conectados en tiempo real a una base de datos para evitar la validación de un mismo ticket más de una vez. No obstante, incluso en ese caso, un usuario legítimo puede ver negada su entrada a un evento si un ticket duplicado sin su conocimiento ha sido utilizado anteriormente. Para resolver este problema, cabe la posibilidad de validar manualmente la identidad del usuario, lo cual supone un proceso lento y costoso.

2 Solución CredoTicket

El sistema patentado CredoTicket permite la generación de documentos digitales consumibles (tickets y cupones electrónicos) y la comprobación off-line de su validez, impidiendo la reutilización, todo ello de forma segura y automática.

La generación de los tickets se realiza mediante un Servidor CredoTicket que utiliza técnicas criptográficas altamente robustas que imposibilitan la creación de tickets válidos por otros medios, lo cual impide la falsificación de tickets. Este servidor genera códigos cifrados que contienen toda la información sobre las características de los tickets. Estos códigos se manipulan como documentos PDF con códigos de barra que el cliente puede imprimir en casa, o como simples códigos numéricos.

Cada ticket generado está asociado a una tarjeta inteligente única, perteneciente al cliente. La comprobación de la validez de los tickets se realiza mediante Lectores CredoTicket que leen el código del ticket y confirman su validez interactuando con la tarjeta inteligente. Para realizar esta validación, el lector no necesita comunicarse con el servidor que generó los tickets (comprobación off-line). Una vez validado, la tarjeta registra el ticket en una lista para no permitir validaciones posteriores, lo cual supone la cancelación efectiva del ticket. Así pues, el sistema asegura que ninguna copia de ese ticket pueda producir fraude, ya que el ticket tan solo puede ser utilizado con una tarjeta determinada y el sistema no permite la reutilización del ticket con la misma tarjeta.

Una misma tarjeta puede realizar un número determinado de validaciones de tickets. Este número depende de la capacidad de la tarjeta. Sin embargo, con tickets con fecha de caducidad, este número puede ser ilimitado.

2.1 Ventajas de CredoTicket

- Sistema seguro y automático que garantiza la autenticidad (impidiendo la falsificación) y la unicidad (impide la reutilización) de los tickets.

- Multitud de medios de acceso utilizados para la obtención de los tickets (Internet,

SMS, etc.), sin que afecte a la seguridad del sistema.

- Para la obtención del ticket, el usuario no precisa de ningún equipo especial en su ordenador, como un lector de tarjetas inteligentes. El usuario sólo requiere ser portador de una tarjeta CredoTicket de bajo coste que utilizará en el momento de la validación o consumo del ticket.

- Permite la validación off-line sin necesidad de conexión a bases de datos o servidores remotos.

- Rápido. Tiempo de validación inferior a 1 segundo.

- Ticket transferible. Si la aplicación lo requiere, permite el anonimato del usuario. El sistema puede ser utilizado por cualquier usuario que sea portador de una tarjeta CredoTicket, pero no requiere de ninguna identificación personal del usuario (la tarjeta puede ser impersonal y transferible).

- Permite múltiples servicios o aplicaciones. Una misma tarjeta se puede utilizar para diversas aplicaciones simultáneamente.

- Permite múltiples operadores o emisores de tarjetas.

- Fácil integración con los sistemas actuales de expedición de tickets.

2.2 Aplicaciones de CredoTicket

- Entradas para eventos (deportes, teatros, cines, etc). Estas entradas pueden dar derecho a varios servicios a la vez. El sistema es particularmente útil en eventos con gran afluencia de público en los que el acceso debe ser rápido y no está controlado por personal.

- Billetes y tarjetas multiviaje de transporte (trenes, metro, autobuses, barcos, aviones, etc.).

- Bonos de hotel, de festivales o similares, en los que ni la fecha ni el destino están prefijados de antemano.

- Cupones de recarga de monederos electrónicos.

- Permisos de acceso únicos, para aplicaciones de control de accesos, como por ejemplo la obtención remota de derechos de acceso.

- Billetes de lotería y apuestas.

2.3 Funcionamiento del sistema

Podemos distinguir las siguientes fases de funcionamiento:

1. Fase de generación de un ticket. El usuario se conecta al servidor del proveedor del servicio (mediante un navegador o móvil) y selecciona las características del ticket deseado, especificando el número de su tarjeta inteligente. El servidor realiza una encriptación 3DES de todos los datos relevantes del ticket con 2 claves que dependen del grupo de lectores del servicio y de la tarjeta del usuario. El usuario recibe un código cifrado resultante, habitualmente en formato de documento PDF con código de barras (permitiendo la impresión en casa del usuario), pero también en otros formatos en función de la aplicación (por ejemplo, mediante códigos numéricos enviados en un mensaje SMS).

2. Fase de validación. El usuario presenta el ticket y la tarjeta al lector, el cual lee el código del ticket y establece una comunicación segura con la tarjeta. La cooperación entre lector y tarjeta permite el descifrado del ticket con las 2 claves: la que depende del lector o servicio y la que depende de la tarjeta.

Si la validación se realiza con éxito (descifrados correctos y los datos son los esperados), la tarjeta registra el ticket en una lista de tal manera que no se permite una posterior validación. Simultáneamente se eliminan de la lista todos los tickets que ya hayan caducado (y por lo tanto, no son válidos independientemente de que estén o no en la lista), consiguiendo de esta manera un alto grado de reutilización de la memoria de tarjeta. Según las aplicaciones, el ticket puede representar el derecho a una determinada cantidad de consumos, que pueden consumirse a lo largo del tiempo (por ejemplo, monederos electrónicos o talonarios de bonos)..

3 Conclusiones

Se ha desarrollado un sistema seguro y muy flexible para la certificación remota de documentos. Se han implementado tanto los servidores necesarios como los validadores a un nivel de prototipo industrial en el que se ha usado microprocesadores de alta seguridad.

Una demostración del sistema servidor puede realizarse en la web del Proyecto CREDO: http://credo.upc.es

Referencias

[1] F.J. Rico-Novella et al., “Method of sending and validating documents”, WO 03/060782, patente de invención.

Distribución de Información Segura con Qos en Entornos Telemáticos (DISQET: TIC2002-00818)

Miguel Soriano, Mónica Aguilar, Oscar Esparza, Marcel Fernández, Jordi Forné, Juan Hernández-Serrano, Isabel Martín, José Luis Muñoz, Esteve Pallarès, Josep Pegueroles, Marcos Postigo
Departament d'Enginyeria Telemàtica. UPC
C/ Jordi Girona 1,3, edificio C3. Campus Nord UPC.
08034 – Barcelona
Teléfono: 934016011 Fax: 93 401 5981
E-mail: {soriano, maguilar, oscar.esparza, marcel, jforne, jserrano, imartin, jose.munoz, esteve, josep, mpostigo}@entel.upc.es

Abstract*. Multicast technology allows efficient bandwidth utilization in emergent multipoint services, such as multivideoconference and video-quasi-on-demand. Security aspects are critical in this scenario. In particular we have to guarantee confidentiality, integrity, authentication of authorised parties, non repudiation and intellectual property rights. On the other hand, multimedia applications require guaranteed network resources to achieve user QoS. New concepts, definitions, protocols, algorithms and mechanisms have been proposed and developed in multimedia applications. This project points two complementary issues: security and QoS. Security aspects studied are: group key management, attribute certificates (for access control) and fingerprinting for intellectual property rights protection. QoS items studied are: the service of elastic, inelastic and semi-elastic flows; and cost analysis depending on the number of the clients that simultaneously access to a resource.*

1 Introducción

El proyecto DISQET aborda básicamente dos problemáticas complementarias vinculadas a la distribución segura de documentos en entornos telemáticos. El primer aspecto es la seguridad y el segundo es la calidad de servicio.

En cuanto a seguridad se refiere, los temas tratados son la gestión de usuarios, evaluando distintas políticas de control de acceso para distintos entornos, protección de la propiedad intelectual y la seguridad cuando se aprovechan las posibilidades del multicast para la distribución de documentos.

Por último, dependiendo de la aplicación, es posible que haya requisitos temporales y al mismo tiempo de calidad de servicio. Para conseguir la QoS requerida por estas aplicaciones, es crucial redefinir problemas como: gestión de los recursos, dimensionado de los nodos intermedios de la red, enrutado interdominio e intradominios, soporte y fiabilidad al transporte

A lo largo del proyecto se han desarrollado soluciones para cada una de estas temáticas. En el siguiente apartado se mencionan las principales contribuciones en el área de seguridad (gestión de usuarios, protección del copyright y seguridad en multicast). En la Sección 3 se destacan los estudios realizados en el ámbito de calidad de servicio y finalmente la Sección 4 presenta las conclusiones del trabajo.

2. Seguridad

2.1 Gestión de usuarios

La gestión de usuarios puede realizarse mediante distintos esquemas, de los cuales el uso de PKI y certificados de identidad y/o atributos es el que goza de mayor aceptación en la actualidad.

En este proyecto se ha abordado una temática fundamental en la PKI: la validación del estado de los certificados, es decir, la capacidad de comprobar en el momento de usar un certificado que éste es válido, es decir, que no está revocado.

En la literatura, los sistemas con mayor aceptación para llevar a cabo dicha validación son las listas de revocación de certificados (CRL) y el protocolo de revocación online OCSP.

En el proyecto se ha desarrollado un testbed (CERVANTES: Certificate Validation Testbed) que permite modelar los distintos sistemas y hacer una evaluación empírica. Asimismo, aprovechando dicha infraestructura, se han propuesto e implementado nuevos esquemas de revocación: H-OCSP [1], ADMHT [2] y E_MHT [3]. Los dos últimos hacen uso de los árboles de Merkle.

En general, los sistemas basados en CRL requieren un mayor ancho de banda que los basados en OCSP (dado que la cantidad de información que ocupa una lista es considerablemente mayor que una respuesta autenticada sobre el estado de un certificado único), pero, en cambio, los basados en OCSP requieren una carga computacional mucho mayor ya que debe

firmar individualmente cada respuesta. Los sistemas basados en árboles de Merkle presentan características intermedias, siendo una buena opción cuando se requieren anchos de banda ostensiblemente menores que en CRL, combinados con una carga computacional en el servidor ostensiblemente menor que en OCSP.

2.2 Protección del copyright

Una de las causas que está frenando la venta de contenidos digitales a través de la red es la dificultad de proteger adecuadamente la propiedad intelectual y los derechos de distribución. Las técnicas que habitualmente se utilizan para resolver estos problemas son respectivamente la de watermarking y la de fingerprinting. La primera consiste en introducir marcas imperceptibles en el documento, de tal forma que identifiquen al autor, y éste pueda demostrar la autoría del documento ante terceras partes. La técnica de fingerprinting, hermana de la anterior también se basa en la inserción imperceptible de datos (marcas), pero la información que se inserta es diferente en cada copia con el fin de poder relacionar cada una de ellas con el receptor correspondiente e identificar si procede al responsable de una distribución fraudulenta.

Las técnicas de fingerprinting involucran los siguientes aspectos:

1. Qué tipo de marca utilizar
2. Algoritmo que permita en tiempo aceptable identificar a los confabuladores a partir de la marca obtenida

El primer aspecto, puede ser cubierto mediante códigos correctores de errores. En tal caso, las palabras código pueden ser usadas como "fingerprinting codewords". Los canales de comunicación no son ideales, de forma que la información recibida puede ser distinta a la transmitida. En el caso que se transmita una palabra código, se precisa un decodificador de canal que estime la palabra que con mayor probabilidad fue enviada. Si el número de errores w es inferior a la mitad de la distancia mínima del código, a la salida del decodificador se obtendrá la única palabra código cuya distancia a la palabra recibida es w. Sin embargo, si el número de símbolos erróneos es mayor, la condición de unicidad no se cumple (es más puede haber alguna palabra código a menor distancia de la recibida que la palabra enviada), de forma que es posible que el decodificador no obtenga la palabra enviada. Por lo tanto, es necesario el uso de técnicas de decodificación complementarias a las empleadas en codificación de canal.

A lo largo de este proyecto se han desarrollado diversas propuestas para decodificación aplicables a esquemas existentes [4], [5], [6], y también se han desarrollado nuevos esquemas completos (codificación y decodificación) [7]. Entre otros resultados, se ha introducido el uso de algoritmos de decodificación con incertidumbre (soft-decoding). La idea fundamental consiste en utilizar técnicas de decodificación indecisa de forma iterativa a partir de los resultados obtenidos en los pasos anteriores; es decir, utilizar el algoritmo tradicional para hallar un conjunto de sospechosos (entre los cuales seguro estaría el primer atacante), y aprovechar esa información para identificar el resto de confabuladores.

2.3 Gestión de claves seguras en multicast

Los servicios multimedia pueden dividirse en dos fases: acceso al servicio e intercambio de datos. La primera fase se realiza normalmente sobre protocolos fiables de transporte y comunicaciones punto a punto. La segunda usa protocolos de transporte no fiable y comunicaciones multicast. Par añadir seguridad a la primera fase basta con aplicar las soluciones unicast existentes. Añadir seguridad sobre IP multicast, en cambio, introduce una nueva problemática no resoluble mediante las técnicas unicast clásicas. A lo largo de este proyecto se han abordado los problemas que aparecen al ofrecer seguridad en multicast y se han propuesto distintas soluciones prácticas.

De todos los posibles ataques en la fase de distribución, la escucha de datos es probablemente el de mayor repercusión. El servicio de protección frente a escuchas es el cifrado. El cifrado multicast añade el problema de gestión de claves de grupos. Para conseguir confidencialidad estricta, la clave de sesión debe actualizarse cada vez que un miembro se da de alta o de baja en el grupo. Cuando el número de miembros es elevado y muy dinámico se hace inviable el reparto punto a punto de claves. En el proyecto se ha propuesto [9] un algoritmo de actualización de claves multicast basado en los árboles lógicos de claves presentados en la literatura, pero que incluye el uso de funciones pseudoaleatorias y propiedades de la teoría de números. Dicha propuesta minimiza la cantidad de memoria necesaria en el Servidor de Claves y reduce considerablemente el ancho de banda requerido por los algoritmos existentes.

En muchos casos la actualización de claves cada vez que un miembro se da de alta o de baja en el grupo no es necesaria. Servicios como la televisión en red, permiten un ligero decremento en seguridad a cambio del ahorro en ancho de banda que supone la espera en la actualización. Ahí operan los algoritmos de renegociación por lotes o en batch que procesan todas las solicitudes de los miembros conjuntamente y de forma periódica. Este tipo de algoritmos deben mantener balanceado el árbol lógico de claves para ser eficientes. Este trabajo propone, implementa y testea un algoritmo en batch que da lugar a árboles lógicos completamente balanceados.

Las técnicas en batch se tratan independientemente al algoritmo de renegociación, pero el estudio conjunto de las técnicas de gestión de claves con estos algoritmos da lugar a mejoras en la eficiencia global.

En el proyecto se ha propuesto la combinación del algoritmo de gestión de claves mencionado anteriormente con las técnicas en batch. El resultado consiste en una mejora en los parámetros de eficiencia. [10]

El algoritmo en batch propuesto, no es soportado por el estándar del IETF de gestión de claves multicast (GDOI). Ello es debido a que nuestro algoritmo permite que los miembros del grupo cambien de posición en el árbol y descarga la responsabilidad de actualización de la posición al mismo cliente. Finalmente, se propone la modificación del protocolo GDOI de forma que se pueda usar con las propuestas de algoritmos de este proyecto [11].

3 Calidad de servicio

Los servicios y aplicaciones ofrecidos en Internet demandan, con mayor frecuencia, un determinado nivel de servicio para su correcto funcionamiento. Esto supone el desarrollo de mecanismos que permitan ofrecer calidad de servicio diferenciada entre extremos distantes de la red. En una red con estas características el coste debido al uso de este servicio diferenciado es mayor que el de utilizar el modo tradicional de transferencia best-effort. Ello obliga a los usuarios a realizar un uso responsable de los recursos reservados. Por tanto, será crucial reservar sólo aquellos recursos estrictamente necesarios para lograr un mínimo coste.

El trabajo desarrollado en esta área tiene una doble vertiente. Por una parte se propone una metodología de construcción de modelos analíticos que permiten evaluar un sistema de Video bajo Demanda (VoD) que adapta su tasa de transmisión a las condiciones de la red [11]. Por otra, se plantea el cálculo analítico de parámetros de nivel de usuario expresados en los acuerdos de nivel de servicio o SLA (*Service Level Agreements*) [12].

Este trabajo tiene como objetivo ofrecer una herramienta para plataformas de gestión dinámica de servicios según el comportamiento de la red, que permita al usuario renegociar su SLA acorde a la demanda de nivel de servicio. Como ejemplo, se ha particularizado dichas expresiones para un servicio de Vídeo bajo Demanda (VoD) que adapta su tasa de salida a las condiciones de red, manteniendo la calidad del flujo de vídeo y variando la tasa del flujo a lo largo del tiempo. El modelo analítico es simple, computacionalmente evaluable y permite obtener medidas *a priori* de rendimiento del sistema. Para ello, se ha modelado el sistema VoD mediante Cadenas de Markov con Recompensas (*Markov-Rewards Chain*, MRC) y en su resolución se ha empleado la técnica de randomización (o uniformización). Los parámetros utilizados han sido obtenidos del análisis estadístico de los flujos de video almacenados en un sistema VoD real. La flexibilidad del modelo permite evaluar medidas de gran interés, por ejemplo, los parámetros de los contratos de nivel de usuario (Service Level Agreement, SLA), y la metodología puede aplicarse a cualquier servicio multimedia caracterizable por un modelo Markoviano, para ello basta asignar un conjunto adecuado de recompensas a la cadena de Markov de tiempo continuo asociada al sistema (*Time-Continuos Markov Chain*, TCMC). Además de varios parámetros SLA básicos utilizados habitualmente, se han propuesto parámetros adicionales de nivel de usuario para tener una concepción extremo a extremo contemplando la visión real del cliente. Los resultados numéricos obtenidos, han sido comparados con resultados experimentales extraídos de la plataforma de laboratorio SSADE *http://ssade.upc.es/*, mostrando un buen ajuste para evaluaciones de valor medio de diversas medidas de evaluación del rendimiento.

Asimismo, en este proyecto se ha abordado la problemática de minimizar los recursos reservados para una transmisión de un flujo semi-elástico [13]. Este tipo de flujo se caracteriza por necesitar la reserva de más o menos recursos en función del estado de la red, por lo que se plantea el diseño de un sistema cliente-servidor capaz de realizar de forma automática las reservas estrictamente necesarias. Para ello, se ha propuesto un mecanismo de control de la ocupación de la memoria del cliente que permite determinar los periodos en que es necesario utilizar un modo de transferencia best-effort (aquellos en los que la ocupación garantiza la disponibilidad de datos en recepción), y los periodos en que se necesita un modo de transferencia con reserva de recursos para garantizar el llenado de dicha memoria. Del análisis de este mecanismo, se deducen expresiones para el coste de transmisión en función de los diversos parámetros que afectan a la ocupación de la memoria del cliente. De estos parámetros destaca especialmente la tasa de llegada de datos, que depende del estado de la red, y el umbral máximo de ocupación que indica cuándo se puede transmitir en modo best-effort. Gracias al correcto dimensionado de este umbral máximo, el cliente es capaz de minimizar los recursos reservados de la red.

4 Conclusiones

Durante este proyecto se han desarrollado distintas propuestas relativas a la seguridad (gestión de usuarios mediante certificados digitales, protección del copyright y gestión de claves en comunicaciones multicast) y a la calidad de servicio en la distribución de documentos digitales por la red. Dichas propuestas han sido expuestas de forma muy somera a lo largo de este artículo y han permitido la finalización de cinco tesis doctorales (M. Postigo, M. Fernández, J. Pegueroles, J. L. Muñoz y O. Esparza)

Agradecimientos

Este trabajo ha podido ser desarrollado gracias a la financiación obtenida mediante el proyecto DISQET: TIC2002-00818

Referencias

[1] Muñoz, Forné, Esparza, Bernabé, Soriano . "Using OCSP to Secure Certificate-Using Transactions in m-commerce". Applied Cryptography and Network Security (ACNS'03). Lecture Notes in Computer Science, vol 2846, 2003. ISSN/ISBN 0302-9743

[2] Muñoz, Forné, Esparza, Soriano. "Certificate revocation system implementation based on the Merkle hash tree". International Journal of Information Security. , vol 2, no 2, 2004. ISSN/ISBN 1615-5262

[3] Muñoz, Forné, Esparza, Soriano. "E-MHT. An Efficient Protocol for Certificate Status Checking".Information Security Applications (WISA). Lecture Notes in Computer Science, vol 2908, 2003. ISSN/ISBN 0302-9743

[4] Fernandez, Soriano. "Identification Algorithms For Sequential Traitor Tracing". Lecture Notes in Computer Science. INDOCRYPT 2004, vol 3348, 2004. ISSN/ISBN 0302-9743

[5] Fernández, Soriano. "Soft-Decision Tracing in Fingerprinted Multimedia Contents". IEEE Multimedia , vol 11, no 2, 2004. ISSN/ISBN 1070-986X

[6] Fernández, Soriano. "Identification of Traitors in Algebraic-Geometric Traceability Codes". IEEE Transactions on Signal Processing. Supplement on Secure Media, vol 52, no 10, 2004. ISSN/ISBN 1053-587X

[7] Fernández, Soriano. "Fingerprinting Concatenated Codes with Efficient Identification".Information Security (ISC'02). Lecture Notes in Computer Science, vol 2433, 2002. ISSN/ISBN 0302-9743

[8] Pegueroles, Rico-Novella. "Enabling Secure Multicast Using a New Java LKH Rekeying Tool".The Third International Conference on Web Engineering (ICWE). Lecture Notes in Computer Science , vol 2722, 2003. ISSN/ISBN 0302-9743

[9] Pegueroles, Wang-Bin, Soriano, Rico-Novella. "Group Rekeying Algorithm using Pseudo-Random Functions and Modular Reduction".Grid and Cooperative Computing (GCC). Lecture Notes in Computer Science, vol 3032, 2004. ISSN/ISBN 0302-9743

[10] J. Pegueroles, J. Hernandez-Serrano, F. Rico-Novella, M. Soriano “Adapting GDOI for balanced batch-LKH” Internet draft. Junio 2003

[11] I. V. Martín F., Juan J. Alins-Delgado, Mónica Aguilar-Igartua, Jorge Mata-Díaz. "Modelling an Adaptive-Rate Video-Streaming Service Using Markov-Rewards Models".First International Conference on Quality of Service in Heterogeneous Wired and Wireless Networks (QShine 2004), 2004. ISBN 0-7695-2233-5

[12] I. V. Martín F., Juan J. Alins-Delgado, Mónica Aguilar-Igartua, Jorge Mata-Díaz. "Analytical Definition of SLA Parameters in a Video-On-Demand Service".12th IEEE International Conference on Networks (ICON2004), Workshop COQODS, 2004. ISBN 0-7803-8783-X

[13] M. Postigo-Boix, Contribución al Estudio y Diseño de Mecanismos Avanzados de Servicio de Flujos Semi-Elásticos en Internet con Garantías de Calidad de Servicio Extremo a Extremo, PhD Thesis. UPC. Abril 2003.

Definición y Desarrollo de Técnicas Basadas en el Conocimiento para su Aplicación a la Gestión de Redes y Servicios: Gestión Semántica (TIC2002-00934)

Jorge E. López de Vergara[†], Víctor A. Villagrá, Julio Berrocal
Departamento de Ingeniería de Sistemas Telemáticos. Universidad Politécnica de Madrid
ETSI de Telecomunicación. Av. Complutense, s/n. 28040 Madrid
E-mail: {jlopez, villagra, berrocal}@dit.upm.es
[†]Actualmente en la Universidad Autónoma de Madrid

Abstract*. The goal of this research project is the improvement of current network management techniques through the study and application of formal ontologies. They would eventually allow the definition of management information in an optimal way, allowing a semantic integration of all the information that currently belongs to different management domains in the same model. In this way, the management information handling could be simplified, avoiding managers' current problems, which have to deal with independent information for each domain, and without the possibility of setting up direct relationships between them.*

1 Introducción y antecedentes

Ante la heterogeneidad de modelos de Gestión de Red Integrada (SNMP, OSI, DMI, WBEM), es muy probable el hecho de que en un mismo dominio de gestión se necesite aplicar varios de estos modelos simultáneamente. Por ello ha sido necesario tradicionalmente establecer mecanismos que posibiliten la interoperabilidad entre los distintos dominios implicados, siendo ésta viable si se consigue definir un conjunto de reglas que traduzcan el protocolo de comunicaciones y modelo de información de gestión. Sin embargo, un aspecto que hasta la fecha no ha tenido fácil solución ha sido cuando en dos dominios distintos de gestión se representaba un mismo concepto de distinta manera: una traducción meramente sintáctica del modelo de información de gestión origen no proporciona el concepto existente en el modelo destino. Se hace por tanto necesaria una traducción semántica que haga corresponder directamente los conceptos de ambos dominios.

En este contexto, este proyecto ha tenido como objetivo general proponer soluciones que permitan perfeccionar la interoperabilidad en lo que se refiere a los modelos de información de gestión, ahondando en las posibilidades que existen para llevar a cabo la traducción semántica. Para ello, la técnica de representación del conocimiento conocida como ontología es una de las respuestas más adecuadas a este problema, dado que proporciona las construcciones necesarias para añadir semántica a la información representada.

La aplicación de esta tecnología al campo de la Gestión de Red ha demostrado ser una de las claves que permite realizar una gestión realmente integrada e inteligente de los diversos recursos que posee una empresa, que normalmente pertenecen a distintos dominios de gestión: Conmutadores y encaminadores gestionados con SNMP, ordenadores con DMI o servicios de comercio electrónico accesibles con una interfaz CORBA se pueden gestionar de manera unificada desde un gestor con un único modelo de información.

Además, el empleo de ontologías también ha permitido la definición de reglas que definen el comportamiento del gestor, de forma añadida al modelo de información de gestión que se haya especificado, definiendo así toda la información relativa a la gestión de manera unificada.

2 Objetivos del proyecto

A partir del objetivo general comentado anteriormente, se ha trabajado en la consecución siguientes objetivos concretos:

1. Formalizar los distintos modelos de información de gestión existentes, capturando aquellas características que debe poseer un modelo de información de gestión de red para que permita trabajar con dicha información a un nivel semántico, basándose en las posibilidades que ofrecen las ontologías.

2. Definir una ontología para la gestión de red. Esta ontología proporcionará la base de información de gestión necesaria para la gestión de redes, sistemas, servicios y aplicaciones, y también modelará y declarará el comportamiento de un gestor de red.

3. Establecer correspondencias entre esta ontología y los modelos de gestión existentes. Esto permitirá a una entidad gestora obtener información de los distintos dominios gestionados desde un punto de vista de información único.

4. Aplicar las reglas de comportamiento definidas a un sistema inteligente de gestión. Esto permitiría comprobar la posibilidad de integrar la totalidad de la información que necesita un gestor en una misma ontología.

5. Experimentar un caso de gestión integrada utilizando las correspondencias previamente definidas para la gestión de un servicio de comercio electrónico. Esto permitirá validar el trabajo realizado, aplicándolo a un caso en el que intervienen varios dominios de gestión: SNMP para la red de datos, WBEM o DMI para el servidor y CORBA para el servicio electrónico.

3 Resultados y conclusiones

Los objetivos planteados en este proyecto se han alcanzado, obteniéndose interesantes resultados entre los que cabe destacar los siguientes:

- Análisis y Comparación de la expresividad semántica de los modelos actuales de información de gestión, según las características que poseen los lenguajes de ontologías. De este estudio se ha concluido que CIM es el lenguaje con mayor expresividad semántica del conjunto analizado.

- Propuesta de formalización de CIM con un conjunto de reglas descritas en OCL, que permita mejorar la capacidad semántica de lenguajes de especificación de información de gestión.

- Propuesta de utilización de un lenguaje de ontologías (OWL, el Lenguaje de Ontologías de la Web) para expresar información de gestión. Establecimiento de reglas de correspondencia entre construcciones de información de gestión y dicho lenguaje.

- Definición del método M&M de fusión y correspondencia para la integración semántica de modelos de información de gestión. Validación de dicho método en un caso de estudio.

- Propuesta para incluir la definición de reglas de comportamiento en la información de gestión, ya sea mediante la inclusión de calificadores en CIM, o mediante SWRL (*Semantic Web Rule Language*, Lenguaje de Reglas de la Web Semántica) en ontologías definidas con OWL.

- Aplicación de estos resultados a la arquitectura de WBEM, en la que el CIMOM maneja el modelo fusionado y usa las reglas de comportamiento, mientras que los proveedores que dan acceso a los distintos dominios de gestión manejan las reglas de correspondencia generadas con el método M&M.

- Desarrollo de distintas herramientas para traducir modelos de información (MIBs de SNMP y CMIP, esquemas CIM) a modelos fusionados basados en lenguajes tradicionales de gestion de red o lenguajes basados en ontologías.

- Desarrollo de herramientas para la generación de ontologías de fusión y correspondencia que permitan la utilización de estos modelos fusionados para acceder a recursos basados en modelos de gestión tradicionales.

Adicionalmente, el trabajo desarrollado en este proyecto y su validación a través de las publicaciones realizadas ha permitido ser invitados a participar en diversas propuestas de proyectos y redes de excelencia del VI Programa Marco de la Unión Europea relacionadas con el tema.

Publicaciones

A continuación se presenta un extracto del conjunto de publicaciones más relevantes que se han realizado en el contexto de este proyecto:

[1] Jorge E. López de Vergara, Víctor A. Villagrá, Julio Berrocal, Applying the Web Ontology Language to management information definitions, IEEE Communications Magazine, special issue on XML Management, Vol. 42, Issue 7, July 2004, pp. 68-74. ISSN 0163-6804

[2] Jorge E. López de Vergara, Víctor A. Villagrá, Julio Berrocal, Benefits of Using Ontologies in the Management of High Speed Networks, Lecture Notes in Computer Science, Vol. 3079, Springer-Verlag. ISSN 0302-9743

[3] Jorge E. López de Vergara, Víctor A. Villagrá, Julio Berrocal, Gestión Semántica: Aplicando las Ontologías a la Gestión de Red, Actas de las IV Jornadas de Ingeniería Telemática (Jitel'03), Gran Canaria, España, 15-17 de septiembre de 2003. ISBN 84-96131-38-6.

[4] Jorge E. López de Vergara, Víctor A. Villagrá, Juan I. Asensio, Julio Berrocal, Ontologies: Giving Semantics to Network Management Models, IEEE Network, special issue on Network Management, Vol. 17, No. 3, May/June 2003. ISSN 0890-8044.

[5] Jorge E. López de Vergara, Víctor A. Villagrá, Julio Berrocal, Juan I. Asensio, Roney Pignaton, Semantic Management: Application of Ontologies for the Integration of Management Information Models, Proceedings of the Eighth IFIP/IEEE International Symposium on Integrated Network Management, Colorado Springs, Colorado, U.S.A. 24-28 March 2003. ISBN 1-4020-7418-2.

ELAS: Elementos Activos de Seguridad

Santiago Felici, Emilio Mira, Jose Duart
Departamento de Informática. Universitat de València
Av/ Vte Andrés Estellés s/n. Campus de Burjassot (Valencia-España)
E-mail: {santiago.felici, jose.m.duart}@uv.es

Abstract. *New trends in network security focus on Instrusion Detection Systems (IDS). But many intrusions cannot be detected (false negative) because the attacks are becoming more and more sophisticated and complex. Then, different kind of IDS are used: based on network or host IDS, based on the misuse detection (signatures) or based on the abnormal detection, with passive or reactive behaviour, etc. As a consequence, a new problem appears, that are the false positive due to the great bulk of information that these systems generate. In the ELAS project, we have faced this situation by the integration of selected open source IDS (Snort or Prelude-NIDS, Bro and Systrace) and the correlation of alerts under the same interface to monitor the whole network. Furthermore, we have researched in Kernel Based Solutions in the Operating Systems, both to improve its security and to analyze how attackers exploit the vulnerabilities.*

1 Introducción

ELAS (Elementos Activos de Seguridad) es un proyecto financiado por el Ministerio de Educación y Ciencia con referencia TIC 2002-04686. El objetivo de este proyecto es velar de forma activa y continua por la seguridad de las redes de ordenadores, en concreto minimizar en lo posible el efecto dañino de los ataques a una red cuando presenta diferentes vulnerabilidades y está ante la amenaza de intrusión.

Además de la seguridad perimetral (*routers*, cortafuegos, etc) diferentes elementos de seguridad se han introducido en una red para evitar la intrusión. Cabe destacar entre ellos redes trampa (*honeynet*), detectores de intrusión (IDS) o de prevención (IPS), evaluadores de vulnerabilidades, etc. De todos ellos, los que más beneficios inmediatos aportan a las redes de producción son los IDS. Sin embargo, dado que el tipo de intrusión puede ser de lo más diversa, desde denegación de servicio, usurpación/alteración de información, suplantación, etc se han diseñado diferentes IDSes en base a diferentes técnicas.

El informe se estructura en base a las tareas realizadas, desde la revisión del estado del arte o punto de partida, pasando por una propuesta de soluciones alternativas y profundizando en la arquitectura propuesta. Finalmente comentamos los mecanismos de difusión.

2 Punto de partida

En el análisis y estudio de las diferentes vulnerabilidades de las redes, servidores y sus servicios, se llegó a la conclusión que las técnicas basadas en IDS son las más recomendables y efectivas para proteger a una red de los diferentes ataques, especialmente de ataques internos.

2.1 Detección de intrusos

Las posibilidades existentes para comprometer a una red son muchas y muy variadas, casi tantas como tipos de IDS. Un resumen de los diferentes tipos de IDS se encuentra en la literatura en [She03]. Los IDS tanto comerciales como públicos, se pueden clasificar brevemente, según la toma de datos (de la red o del *host*, NIDS o HIDS), según el método de detección (basados en firmas o basados en mal comportamiento) y según su respuesta (activa o pasiva).

El objetivo de un IDS es detectar el máximo de intrusiones posibles, pero como hemos dicho, según su implementación, hay muchas intrusiones que no son capaces de ello. Por tanto, concluimos que una "*red de IDSes*" es la mejor opción para velar por la seguridad, de forma que sea sensible a diferentes ataques. Obviamente, ello no quita que la red tenga una buena política de seguridad, punto clave y piedra angular.

2.2 Normalización de la información de alertas

Uno de los primeros pasos para poder gestionar toda la información procedente de los diferentes IDS, es ponerla en un formato común. Aunque no sea de carácter obligatorio, la mayoría de desarrolladores de IDSes tratan de normalizar este proceso, tanto la forma en cómo representar las diferentes alertas según el formato IDMEF [IDMEF03], como la forma de nombrar los diferentes tipos de ataques, *Common Vulnerabilities and Exposures* [CVE03].

3 Arquitectura de integración de IDS propuesta

El objetivo de esta sección es detallar y justificar la arquitectura utilizada. En particular, en ELAS se ha

optado a disponer de una red de sensores, con características complementarias, evitando así que tras una correlación de su información, se minimicen los falsos positivos y por otro lado, con la diversidad de IDSes no existan falsos negativos; una decisión de compromiso entre fiabilidad y seguridad.

3.1 Elección de IDSes: justificación

Durante la última década se han desarrollado diferentes herramientas para detección de intrusos de libre distribución, sin embargo, la situación actual no está ni mucho menos resuelta. Y es más, el problema se agrava por la gran diversificación existente. Esto ha dado pie a que la reutilización de IDSes de libre distribución para nuestra arquitectura deba cumplir las siguientes características: *rapidez, diferentes técnicas de detección para correlar la información, generen pocos falsos positivos, no tengan falsos negativos y sus salidas puedan ser normalizadas en CVE definidas en formato IDMEF.*

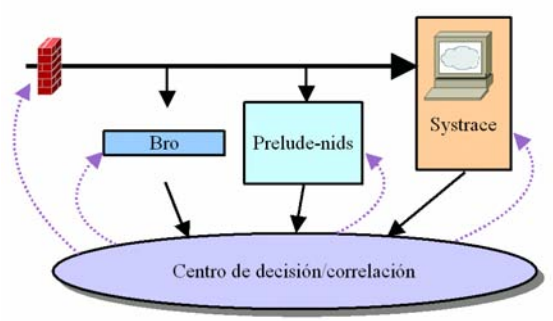


Figura 1: Arquitectura de la red de sensores IDS: detalle de los diferentes elementos integrados en el proyecto ELAS, especificando el flujo de información

Con este criterio la búsqueda se ha centrado por tanto en escoger un NIDS, concretamente *SNORT* (*http://www.snort.org*), un IDS de protocolo, concretamente *BRO* (*http://bro-ids.org*) y un IDS basado en *host* para el servidor que controle las llamadas al sistema, concretamente *Systrace* [Ni03]. Estos sensores se muestran en la figura 1, con un centro de decisión/correlación.

3.2 Integración de sensores

La información generada por los sensores como se muestra en la figura 1 es integrada por un administrador, que a su vez puede integrar funciones de correlación de alarmas/eventos y decisión. Para realizar este proceso de integración, hemos utilizado también herramientas de libre distribución como *Prelude* (*http://www.prelude-ids.org/*). Tras la gestión de la información y según información de los diferentes equipos y servicios instalados, se pueden definir reglas que minimicen los falsos positivos. Por ejemplo, no es una alerta un ataque de "Gusano Rojo" a un servidor Linux Apache.

Una foto de la maqueta de trabajo podemos ver en la figura 2.

Figura 2: Fotografía de la maqueta de pruebas del proyecto ELAS

3.3 Mejora de los núcleos de los servidores

También hemos realizado pruebas y estudios sobre protecciones en los núcleos para evitar vulnerabilidades de los servicios y/o aplicaciones, concretamente "pilas no ejecutables", PaX, etc

4 Difusión de resultados, publicaciones y tesis

En la web del proyecto http://elas.uv.es se distribuyen los informes del proyecto. Algunos de ellos, han sido enviados a revistas con índice de impacto. Además, en la actualidad existen varios PFC y una tesis asociada, basada en mecanismos eficientes de correlación de IDS diferentes.

Referencias

[CVE03] Common Vulnerabilities and Exposures, http://www.cve.mitre.org/, 2003.

[Den87] D.E. Denning, "An Intrusion Detection Model," IEEE Trans. Software Eng., vol. 13, no. 2, pp. 222-232, Feb. 1987.

[IDMEF03] D. Curry and H. Debar, "Intrusion Detection Message Exchange Format: Extensible Markup Language (XML) Document Type Definition," IETF draft, Jan. 2003.

[Ni03] Niels Provos, "Improving Host Security with System Call Policies", 12th USENIX Security Symposium, Washington, DC, August 2003.

[She03] J.Sherif, R. Ayers and T.G. Dearmond, "Intrusion detection: the art and the practice. Part I & II", Information Management and Computer Security, 2003.

Red de Acceso Celular IP Multisalto (RACIMUS)

R. Sanz, L. Muñoz, M. García, J.A. Irastorza
Grupo de Ingeniería Telemática
Departamento de Ingeniería de Comunicaciones. Universidad de Cantabria
Av. Los Castros S/N. 39005 – Santander
E-mail: [roberto, luis, marta, angel]@tlmat.unican.es

***Abstract**. Wireless will be playing an increasingly important role in the future Internet. RACIMUS studies how multi-hop heterogeneous wireless networks (MHWN) can support mobility of users, packet routing and adaptation to varying link conditions. We aim to specify, implement, simulate and demonstrate a heterogeneous multi-hop wireless IP network consisting of several wireless technologies. We will be providing end-to-end optimization for IPv4 and v6 based services over MHWN with respect to throughput, power consumption and implementation complexity. The results from previous European projects using a platform independent Performance Enhancing Proxy (PEP) will be exploited, streamlined and extended. RACIMUS does not aim to reach the same goals as in MANET by considering semi ad-hoc networks taking real device constraints into account.*

1 Introducción

El rápido desarrollo de los servicios en Internet y la evolución exponencial de su número de usuarios ha coincidido en el tiempo con la eclosión de la telefonía móvil celular de segunda y tercera generación, dibujándose un escenario en el que la convergencia entre Internet y las comunicaciones inalámbricas sustentarán las bases de una nueva forma de relacionarse la sociedad con su entorno. El mayor desafío tecnológico es la unión de los sistemas móviles con Internet, lo que representa claramente el presente y el futuro de la base para las aplicaciones de la Sociedad de la Información [1]. Esto requerirá una combinación de las especificaciones de sistemas móviles del 3GPP (3rd Generation Partnership Project), IP/Multimedia y otros grupos pertenecientes al IETF (Internet Engineering Task Force). El grupo de trabajo del IETF denominado MANET (Mobile Ad hoc Networks) está trabajando en la elaboración de estándares relacionados con este campo, aunque éste plantea un entorno más restrictivo que el proyecto RACIMUS. Actualmente, existen diferentes tipos de accesos inalámbricos: fijo (FWA, Fixed Wireless Access), de área local (WLAN, Wireless Local Area Network), celulares (GSM/UTRA) y personales (WPAN, Wireless Personal Area Network). El resultado final consiste en un gran conjunto de diferentes estándares y protocolos incompatibles entre sí, lo cual provoca que el usuario final disponga de un variado conjunto de dispositivos conectados a diferentes redes de acceso.

1.1 Objetivos

El objetivo de RACIMUS consiste en especificar, diseñar, desarrollar, demostrar y validar una red inalámbrica multisalto heterogénea (MHWN) basada en la pila de protocolos TCP/IP, consistente en la combinación de tecnologías WPAN, WLAN y FWA (ver Fig. 1). El objetivo final es la optimización global entre extremos de los servicios considerando throughput, consumo y complejidad de implementación. El trabajo realizado en RACIMUS es estratégico para comprender y extender la interoperabilidad entre IPv6 y un entorno inalámbrico bajo demanda. A su vez, se trata de validar la interconectividad entre las islas inalámbricas IPv6 y la infraestructura IPv4. Aunque RACIMUS se basa en IPv6, se proporcionan los mecanismos para la transición y la interoperabilidad con redes IPv4. Por otro lado, puesto que se proporcionan servicios IPv6 independientes de la capa física de cada plataforma, se evalúan las nuevas soluciones relativas a infraestructuras radio nacidas de recientes proyectos IST (CABSINET, WINDFLEX) en los que el GIT (Grupo de Ingeniería Telemática) ha trabajado activamente. Finalmente, se contempla la posibilidad de colaborar con proyectos en los que se trabaje sobre la QoS (Quality of Service), pues las soluciones extraídas de RACIMUS pueden utilizarse de manera transparente para evaluar diferentes métodos de QoS.

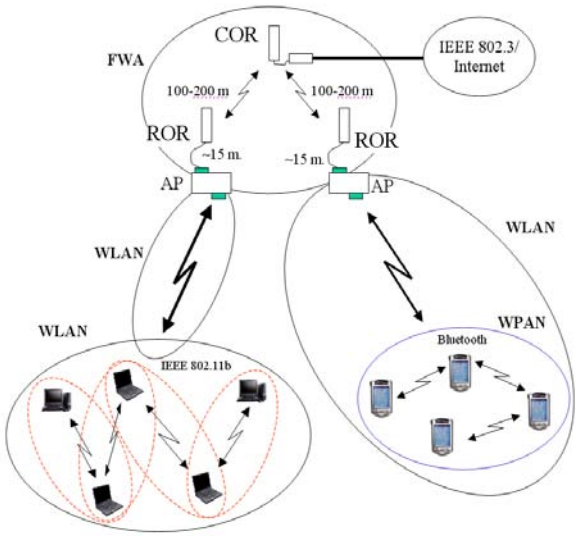


Figura 1. Esquema de la red propuesta en RACIMUS

2 Tecnologías de acceso

Actualmente, los terminales móviles poseen la capacidad de incorporar funcionalidad WPAN mediante Bluetooth haciendo posible la interconexión de muchos dispositivos en pequeñas células. Al mismo tiempo, los terminales están evolucionando para soportar la modalidad de multi-modo multi-estándar, para que un mismo dispositivo pueda conmutar su transceptor pasando de Bluetooth a IEEE 802.11 o viceversa, con objeto de conectarse a un punto de acceso (AP) de una WLAN (handover vertical). Otros ejemplos serían la conexión de varias WPANs en una sala de conferencias mediante configuración ad hoc, o la utilización de varios APs WPAN o WLAN en transportes públicos para acceder a la red fija a través de sistemas celulares 2.5/3G.

3 Protocolos de encaminamiento

La información suministrada por entidades de nivel de enlace, transparentes a las pilas de protocolos de capas superiores, tradicionalmente conocidas como PEPs (Performance Enhancing Proxies), como por ejemplo la tasa de paquetes erróneos, medidas de calidad del enlace, etc, proporcionan una métrica adicional en la que se basarán dichos algoritmos para realizar las decisiones de encaminamiento eficientemente (ver Fig. 2). Se ha elegido el protocolo reactivo DSR (Dynamic Source Routing) dado que permite optimizar de forma multiparamétrica las prestaciones extremo a extremo para los distintos flujos de datos [2]. Adicionalmente, se le ha añadido el mecanismo de HELLO propio de un protocolo preventivo como es AODV (Ad hoc On demand Distant Vector), representando un paso adelante respecto al estado del arte de los protocolos de encaminamiento.

4 Resultados

Se han diseñado e implementado las funcionalidades que permiten la utilización, por parte de los nodos participantes en la red multisalto, de aquellos módulos que permiten mejorar las prestaciones de las sesiones soportadas por las distintas interfaces radio.

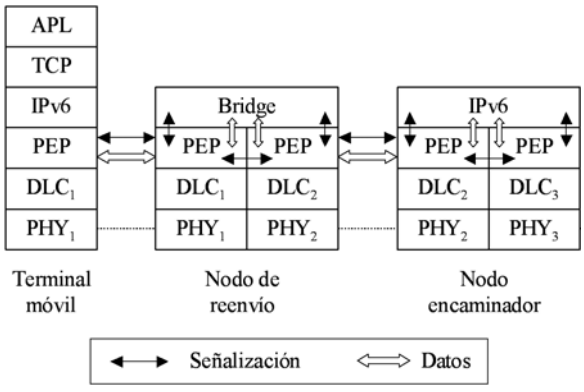


Figura 2. Pila de protocolos de RACIMUS

Así, se ha implementado una funcionalidad destinada a descubrir las capacidades de un nodo adyacente para, de este modo, poder acordar la utilización de un módulo específico adaptado a las características del enlace. Para dar soporte a dicha funcionalidad, se ha concebido un protocolo de señalización basado en el intercambio de datagramas UDP entre las diferentes entidades PEP. Asimismo, se han realizado pruebas de campo para determinar las prestaciones de una red IEEE 802.11b tanto para un único salto como para una red ad hoc multisalto [3].

Se ha realizado el diseño y la implementación de la señalización intra-nodo PEP-to-PEP, del PEP forwarding, del tunneling entre puntos de acceso así como de la fragmentación. Se han diseñado e implementado las técnicas de movilidad en MHWN, realizando las últimas pruebas de funcionamiento tanto de los protocolos de encaminamiento como de las técnicas de movilidad en redes ad hoc antes de proceder a la integración final de todos los módulos programados que tendrá lugar durante el último semestre del proyecto.

5 Conclusiones

En el marco del proyecto RACIMUS se han definido algunos de los posibles mecanismos y arquitecturas de referencia para implantar de manera transparente redes coexistentes multioperador; se han diseñado redes en las que coexisten islas ad hoc, sistemas inalámbricos trabajando en modo infraestructura con conectividad a redes cableadas, inalámbricas fijas y celulares; se ha propuesto un esquema de migración rápida hacia IPv6 y se han concebido mecanismos transparentes que evitan modificar los protocolos de nivel de red y transporte de la red inalámbrica heterogénea.

Agradecimientos

Agradecer el soporte proporcionado por la División Comunicaciones Inalámbricas de Fagor Electrónica ubicada en el Centro de Desarrollo Tecnológico de la Universidad de Cantabria (CDTUC) en la aplicación de los resultados del proyecto.

Referencias

[1] C.E. Perkins, "Mobile IP", IEEE Comm. Magazine, vol. 35, no. 5, pp. 84-99, Mayo 1997.

[2] R. Agüero, L. Sánchez, J. Choque, R. Sanz, L. Muñoz, J.A. Irastorza, “On the implementation and experimental characterization of the Dynamic Source Routing protocol for mobile ad hoc networks”, Wireless Pers. Multimedia Comm. (WPMC), Oct. 2003, Yokosuka, Japón.

[3] L. Muñoz, R. Agüero, J. Choque, J.A. Irastorza, L. Sánchez, M. Petrova and P. Mähönen, "Empowering Next-Generation Wireless Personal Communication Networks", IEEE Comm. Magazine, pp. 64-70, Mayo 2004.

Gestión **I**nteligente de los **R**ecursos Radio para Redes **A**d-hoc de Alta Velocidad a Través del Desarrollo de Técnicas de: Lógica **F**uzzy, Procesado de Señal y de Protocolos de Control de **A**cceso al Medio[1]

Ana I. Pérez Neira, Antonio Pascual Iserte, Montserrat Nájar
Universidad Politécnica de Cataluña - UPC
c/ Jordi Girona 1-3, módulo D5-201, 08034- Barcelona.
Teléfono: 932046436. Fax: 932046447
E-mail: anuska@gps.tsc.upc.edu

Abstract*. The general goal of Girafa is to increase throughput and quality of service (QoS) in ad-hoc networks by applying strategies of reconfigurability and PHY/MAC cross-layer design and by developing techniques involving fuzzy logic and advanced signal processing.*

1 Introducción

El propósito del proyecto es abordar diferentes problemáticas que pueden darse en redes ad-hoc y estudiar posibles alternativas para dotarlas de alta velocidad de transmisión y calidad de servicio. El proyecto parte de herramientas disponibles en la capa física para enfrentarse al medio radio y estudia cómo la capa de acceso puede interactuar con la capa física para conseguir mejoras en el acceso y en ofrecer una calidad de servicio integrada. El proyecto se divide en siete actividades/tareas principales de investigación: Tarea 1 (T1) o desarrollo de sistemas reconfigurables a nivel físico; Tarea 2 (T2) o desarrollo de sistemas reconfigurables a nivel de enlace/sistema; Tarea 3 (T3) o estrategias de optimización conjunta (PHY-MAC); Tarea 4 (T4) o fusión de las técnicas desarrolladas a nivel de MAC en T2 y T3; Tarea 5 (T5) o desarrollo de sistemas de lógica fuzzy y de procesado avanzado de la señal para gestión del nivel de red; Tarea 6 (T6) o desarrollo de una plataforma; y Tarea 7 (T7) o estudio de arquitecturas hardware disponibles en el mercado.

A continuación se pasa a describir el trabajo desarrollado en cada una de las tareas por parte de la UPC. En www.cttc.es/projects/girafa está disponible una página web que divulga los resultados del proyecto coordinado. Entre otros se puede encontrar [1], que contiene las publicaciones que se han realizado en cada una de las tareas.

2 Actividad realizada y resultados

El trabajo a nivel de capa física que se ha realizado en la tarea T1 se ha centrado en los sistemas WLAN (*wireless local area network)*, concretamente en los estándares IEEE802.11 a, b y g, o modificaciones de los mismos para redes de seguridad pública, que operan en la banda de 5 GHz (a) y 2.4 GHz (b y g). Principalmente se ha considerado OFDM como medio de transporte; ya que se muestra como una buena alternativa para combatir los ecos del canal, incluso para los nuevos estándares móviles de banda ancha, 802.16 y 802.20. En esta tarea se han diseñado sistemas de transmisión punto a punto, siendo la transmisión multipunto objeto de otras tareas del proyecto. Para resolver el problema de interferencias y ganar en QoS se han desarrollado tanto técnicas de procesado de arrays como técnicas de lógica fuzzy. Por otra parte, el ofrecer arquitecturas reconfigurables según el canal o escalables y robustas según la información disponible del canal de transmisión ha sido un común denominador en T1.

Para mejorar la capacidad y comportamiento de los sistemas multiantena existentes, las técnicas de procesado de array desarrolladas se han centrado en cómo emplear de manera inteligente múltiples antenas en transmisión en sistemas MIMO (*multiple input multiple output*). Concretamente, se han tenido en cuenta las limitaciones por tener una CSI o *channel state information* imperfecta, más crítico en transmisión que en recepción. Los diseños van dirigidos a conseguir una mejora de diversidad y BER (*bit error rate*), o de capacidad y velocidad de transmisión y estudiar estrategias para dotar al transmisor de capacidad de reconfigurabilidad o adaptación al enlace. Una herramienta importante en los estudios realizados ha sido la optimización convexa, que ha permitido obtener soluciones de diseño generales para el procesado MIMO. Otras herramientas matemáticas útiles que se han aplicado y desarrollado para el diseño de sistemas multi-antena en recepción prácticos o implementables ha sido el procesado fuzzy y el análisis matricial asintótico.

El procesado fuzzy también se ha empleado por la UPC para desarrollar receptores de espectro ensanchado por código escalables a diferentes tipos

[1] Contrato: TIC2002-04594-C02-01.

de interferencia pues son capaces de enfrentarse no sólo a interferencias de banda estrecha, sino también de banda ancha. El objetivo ha sido diseñar receptores que, sin necesidad de diversidad espacial, sean capaces de trabajar en la banda de 2.4 GHz, que, al ser libre de licencia, no ofrece un entorno radio libre de interferencias a los sistemas que trabajan en ella. Una vez más, las técnicas de lógica fuzzy han permitido obtener sistemas de sencilla implementación y fácil ajuste.

En cuanto al tipo de red WLAN, considerado en T2 y T3, hablamos tanto de redes de acceso centralizado, como descentralizado, si bien el primero ha sido el más considerado a la hora de garantizar QoS y elevada capacidad. La tarea T2 se ha dividido en dos objetivos: i) el diseño de sistemas de enlace reconfigurables, desarrollada principalmente por el CTTC; ii) el diseño de *schedulers* de capa física capaces de gestionar los diferentes recursos de nivel físico: potencia, frecuencias, espacio y arquitecturas de transmisión. Tal y como era previsto, en ambos objetivos, la Tarea 2 va convergiendo con la Tarea 3. Esta última está centrada en técnicas intercapa o *cross-layer* PHY-MAC para el diseño de protocolos y, por lo tanto, tiene una visión más global de la red y del sistema de comunicaciones. Prueba de ello es que, por ejemplo, resultados de los estudios *cross-layer* nos permiten concluir que las técnicas que manejan *goodput* como medida de calidad que tiene en cuenta tanto la BER como la velocidad de transmisión, serán más adecuadas a la hora de hacer adaptación al enlace. Otro ejemplo es la importancia del tamaño de las colas y no sólo los parámetros del canal si se quiere hacer una gestión de recursos físicos reconfigurable. Y, finalmente, diremos que, para ganar en *throughput* y QoS, hemos combinado sistemas reconfigurables tanto a nivel físico como de acceso (T2), con sistemas en los que el acceso es multiusuario (T3). La tarea T4 es la encargada de compendiar las diferentes técnicas y estrategias desarrolladas en T2 y T3 y estudiar lo conveniente de emplear lógica fuzzy a la hora de reconfigurar el sistema tanto a nivel PHY como MAC. Con dicho objetivo la UPC ha estudiado de manera exhaustiva algoritmos de clasificación fuzzy que permitan reconocer un escenario incorporando conocimiento experto en el caso de que sea disponible. En esta tarea se ha puesto especial cuidado en el grado de abstracción que de la capa física hace la capa MAC para poder reaccionar adecuadamente a los recursos del canal radio. También es importante el diseño de un modelo matemático para protocolos de múltiple acceso así como controlar cuán óptimos son los protocolos de gestión de recursos físicos que se diseñan.

El objetivo principal de la Tarea 5 es ofrecer una calidad de servicio integrada entre capas. Los trabajos realizados por la UPC van en tres direcciones. Por una parte, integrar lo realizado a nivel de MAC-PHY con los requerimientos de la capa de red. Por otra parte, estudiar el hand-off en redes heterogéneas mediante sistemas de lógica fuzzy que sepa garantizar una mínima QoS. Finalmente, se han diseñado estrategias de encaminamiento en redes adhoc, donde las métricas usadas tienen en cuenta explícitamente parámetros relacionados con las baterías disponibles de los nodos y las potencias de transmisión. Se han mejorado métricas ya existentes en la literatura, dando como resultado soluciones que son capaces de incrementar de forma global la vida de la red adhoc . Tanto para el hand-off como para el encaminamiento, es importante conocer la posición de los distintos nodos. Técnicas de procesado avanzado de señal como algoritmos de estimación de alta resolución y algoritmos adaptativos se han aplicado en el desarrollo de técnicas de localización, proporcionando una gran precisión en la estimación de la posición.

En cuanto a la Tarea 6, la UPC ha desarrollado los modelos de canal físico y las interfícies con el simulador de sistema Opnet tanto para SISO (*single input multiple output*), SIMO (*multiple input multiple output*) y MIMO (*multiple input multiple output*). Actualmente se están desarrollando las interfícies para MIMO. Finalmente, en el caso de que sea necesario se estudiarán posibles modos de implementación *hardware* del controlador fuzzy.

3 Colaboración con diversos sectores

A raíz de este proyecto y de los resultados obtenidos se han elaborado varias propuestas en programas europeos. Las concedidas han sido: Marquis (procesado multiantena), Newcom (red de excelencia de procesado avanzado de señal), Widens (redes ad-hoc para seguridad pública), Planets (redes ad-hoc domésticas). En relación con los sistemas de clasificación fuzzy se ha llevado a cabo la Acción Integrada HF2001-55 y un convenio con la empresa Salvio Busquets S.A.

4 Conclusiones

El proyecto se está realizando según lo previsto. Su continuación consistirá en seguir con el cronograma indicado en la propuesta: i) convergencia de T2 y T3 para obtener diseños PHY-MAC lo más óptimos posibles para sistemas WLAN; ii) QoS integrada; iii) finalización del simulador PHY-MAC "software". Observamos también que las estrategias de reconfigurabilidad y cross-layer que se plantearon en este proyecto se están consolidando tanto en las redes de sensores como en los futuros sistemas radio denominados de 4G, que parecen hacer una fuerte apuesta por *cognitive radio*.

Referencias

[1] Informe de seguimiento primera y segunda anualidad UPC, www.cttc.es/projects/girafa/research.htm

Gestión Inteligente de los Recursos Radio para Redes Ad-hoc de Alta Velocidad a Través del Desarrollo de Técnicas de: Lógica Fuzzy, Procesado de Señal y de Protocolos de Control de Acceso al Medio[1]

Miguel Ángel Lagunas Hernández, Marc Realp, Carlos Bader.
Centre Tecnològic de Telecomunicacions de Catalunya -CTTC
Edificio Nexus I, c/ Gran Capità 2-4, planta 2, desp. 202-203. 08034- Barcelona.
Teléfono: 932058415. Fax: 932058399
E-mail: m.a.lagunas@cttc.es

Abstract*. The project mainly encompass researches on radio resource management (RRM) in ad-hoc networks. The introduction of some degrees of knowledge between functionalities in different OSIs layers (i.e. Cross-Layer information, CL) facilitate the overall design of the physical layer, the medium access control, and the RRM. On the other hand, strategies for CL design based on signal processing procedures and Fuzzy logics, constitute an instrument for some strategies considered essential for RRM, namely as the scheduling, robustness exchange in space-time processing, hybrid access mechanisms, etc. During this project some technical responses of the feasibility of the cross-layer concept have been brought and developed to achieve the intelligent RRM in ad-hoc networks. Based on CL, different approaches have been adopted and developed using reconfigurable and adaptable mechanisms.. Obtained results have allowed a rich collaboration in/with Inter/National projects*.

1 Introducción

El objetivo principal del proyecto es como bien lo indica su titulo, la "**G**estión **I**nteligente de los Recursos Radio para **R**edes **A**d-hoc de Alta Velocidad a través del Desarrollo de Técnicas de Lógica: **F**uzzy, Procesado de Señal y Protocolos de Control de **A**cceso al Medio **-GIRAFA**" [1]. Las redes ad-hoc son redes para voz y, sobre todo, para datos que se caracterizan por su flexibilidad para establecer comunicaciones de forma dinámica. Este concepto resulta muy atractivo para comunicaciones WLAN que no cuentan con una infraestructura previa pero si con un coste cada vez inferior, por este motivo su demanda es cada vez mayor y sus requerimientos en cuanto a capacidad son crecientes. No obstante, los sistemas ad-hoc actuales únicamente ofrecen la conexión entre diferentes terminales móviles, pero no están concebidos ni preparados para ofrecer grandes capacidades (*aplicaciones multimedia*) Son redes cuyas prestaciones vienen limitadas principalmente por las interferencias tanto de la propia red como de otros sistemas.

El presente proyecto aborda dicho problema de interferencias de forma novedosa a través de una gestión inteligente de los recursos radio (RRM). Por gestión inteligente se entiende: aumentar la capacidad y la calidad (QoS) de las redes ad-hoc a través de estrategias de reconfigurabilidad y de optimización conjunta entre diferentes capas OSI, y recurrir tanto a técnicas clásicas de procesado de señal como a técnicas de lógica Fuzzy. El proyecto en si se divide en siete actividades/tareas principales de investigación descritas en [1], en las cuales el CTTC participa en: tarea 2 (T2): desarrollo de sistemas reconfigurables a nivel de enlace, tarea 3 (T3): estrategias de optimización conjunta (PHY-MAC), tarea 5 (T5): gestión de recursos para QoS, y en la tarea 6 (T6): desarrollo de una plataforma software.

2. Actividad realizada y resultados

Como primer etapa en T2, se abordaron las principales limitaciones en la reconfigurabilidad durante el acceso, para el cual se hizo un hincapiés en las diversas técnicas multiusuario que reduzcan el nivel de interferencia tanto en el receptor como en el transmisor. La utilización de múltiples antenas permite una mejora de las prestaciones de los sistemas de comunicación inalámbricas. Además, la introducción de técnicas de adaptación a nivel de enlaces, como la modulación adaptativa permite alcanzar una mayor eficiencia espectral. Combinando los dos principios, se ha conseguido el desarrollo de algoritmos de adaptabilidad a nivel de enlace para sistemas de tipo MISO (*con modulación adaptativa*) considerando un canal de retorno para la información de canal con una tasa de bit reducida. Los resultado obtenidos demuestran una mejora significativa a nivel de enlace, del "throughput" manteniendo la potencia de transmisión y la tasa de error de bits por debajo del umbral permitido a cada usuario (ref. 16 en [2]). También, el desarrollo de un procedimiento de acceso diferente, ha permito la reconfigurabilidad de un sistema CDMA con acceso aleatorio Aloha en otro con acceso híbrido TDMA-CDMA, ha permitido que el acceso sea un proceso reconfigurable, adaptándose de forma dinámica a las

[1] Contrato: TIC2002-04594-C02-02.

características de la red (*niveles de trafico, etc*). Los resultados obtenidos han permitido mostrar que el alcance de las prestaciones óptimas depende en gran parte de la carga del tráfico de la red, de la capacidad de los diferentes procedimientos de control de acceso, y de los enlaces de reconfiguración en los diferentes niveles de tráfico dentro de la propia red (ref. 3 y 6 en [2]). Conseguir adaptar la velocidad de transmisión a las condiciones del medio basándonos en estrategias *CL* (ref. 17 en [2]), ha sido también una de las líneas exploradas. Se ha diseñado un esquema de selección de antenas de transmisión, con una velocidad de transmisión por antena constante. Para alcanzar un compromiso entre fiabilidad y velocidad de transmisión se ha propuesto optimizar la expresión del "throughput" en la capa de enlace, observándose considerables mejoras respecto a sistemas donde la interacción entre capas es inexistente (ref. 10, 11 y 12 en [2]). Los avances conseguidos en T2 han permitido una estrecha/activa colaboración en el proyecto Europeo **ANWIRE (red de excelencia)** [3], principalmente en la definición de los diferentes soportes /elementos esenciales (*hardware/software*) de reconfigurabilidad/ adaptabilidad dentro del concepto **ABC** (*Always Best Connected*) en redes ad-hoc (ref. 3 en [2]). Por otro lado, parte de los resultados obtenidos han dado lugar a la elaboración por parte del CTTC de propuestas de investigación en el campo de "**Cognitive Radio**", concepto que se basa en gran medida sobre estrategias de sondeo al medio y de reconfigurabilidad para las futura redes de comunicación.

El primer paso hacia el desarrollo de estrategias de optimización conjunta PHY-MAC en T3, ha sido la evaluación del impacto en términos de "throughput" que pueda haber entre un sistema con MAC basado en el criterio de maximización del "throughput" *(TMC*) y una estrategia basada sobre el criterio de maximización de la capacidad (*CMC*) a nivel de la capa física para las cuales se requieren máximos niveles de BER (*o SNIR mínima*). También ha sido posible el desarrollado de un algoritmo MAC multi-paquete que maximice el "throughput" calculando el número óptimo de transmisiones simultaneas (ref. 5, 13 y 15 en [2]). Además, se ha evaluado el impacto sobre el "throughput" al utilizar un MAC con estrategia de "scheduling" basada en el criterio TMC. Los resultados obtenidos han sido contrastados frente a un PHY con estrategias de "scheduling" basadas en *CMC* (ref. 20 en [2]).

Con la participación de la UPC, en esta ultima fase del proyecto en T5, se colabora en el desarrollo de algoritmos reconfigurables al entorno que permitan reconfigurar la QoS (ref. 21 en [2]) sin necesidad de interrumpir la comunicación y la optimización de los recursos de forma conjunta entre las capas PHY y de enlace. Para la correcta evaluación de las técnicas que se diseñen se ha escogido a la herramienta OPNET Simulator. La primera fase de desarrollo en T6 fue el diseño un sistemas *CDMA* con recepción MPR. Para conseguirlo fue necesario introducir varias modificaciones en el *pipeline* de los transceptores radio que usa la propia herramienta, como el calculo de la SNR, o la capacidad de recepción MPR. En esta ultima fase del proyecto, se está procediendo a la introducción de las estrategias *CL* desarrolladas. La experiencia adquirida en esta tarea sirve actualmente al CTTC para la elaboración de una plataforma software más compleja, capaz de abarcar/evaluar conceptos de gestión adaptativa de los recursos radio a nivel de sistema para sistemas multi-portadoras.

3. Colaboración con diversos sectores

Ha raíz de este proyecto y de los resultados obtenidos, se han elaborado varias propuestas en programas europeos, los cuales destacamos; i) **ANWIRE** [3] ii) **NEWCOM (red de excelencia)** [3] donde el CTTC participa en estrategias de optimización *CL*. iii) El proyecto **MARQUIS** [3] donde destacamos en sus líneas de investigación temas abordados en GIRAFA cuales destacamos; el diseño de capas PHY/MAC con soporte para *QoS*, y el "scheduling" espacial para MACs multi-paquetes. También durante proyecto se establecieron contactos con el centro de investigación **FTW** (*Forschungszentrum Telekommunikation Wien*) de Viena, y con el **grupo de telecomunicaciones AUNA**, para exponer los avances adquiridos por el CTTC en la gestión de los recursos radio de las redes ad-hoc.

4. Conclusiones

El flujo de información intercambiada entre la capa PHY y MAC para la gestión de los recursos radio, es considerada como la primera base al diseño de estrategias "cross-layer". Durante todo el periodo del proyecto se han desarrollado varios algoritmos de adaptabilidad/reconfigurabilidad, se han obtenido abundantes resultados [2] en los cuales se demuestran mejoras en las prestaciones al introducir los cambios novedosos sugeridos frente a los mecanismos existentes en las redes ad-hoc actuales. Gran parte de los planteamientos/estrategias de reconfigurabilidad y de *CL* (PHY-MAC) aportados en este proyecto se están reforzando como conceptos imprescindibles en los futuros sistemas/redes de comunicación.

Agradecimientos

Los autores deseamos agradecer la labor participativa en las distintas tareas de investigación de este proyecto a Miquel Payaró, José López-Vicario, Carlos Antón.

Referencias

[1] http://www.cttc.es/projects/girafa/wp.htm

[2] http://www.cttc.es/projects/girafa/dissemination.htm

[3] http://www.cttc.es/projects/

Estudio y despliegue de movilidad en el entorno de red heterogéneo y multiproveedor del proyecto SAM[1]

Juan Quemada, Tomás P. de Miguel, Tomás Robles
Dpto. Ingeniería de Sistemas Telemáticos de la Universidad Politécnica de Madrid
ETSI Telecomunicación. Ciudad Universitaria s/n Madrid
Teléfono: 915495700 Fax: 913367333
E-mail: {jquemada,tmiguel,trobles}@dit.upm.es

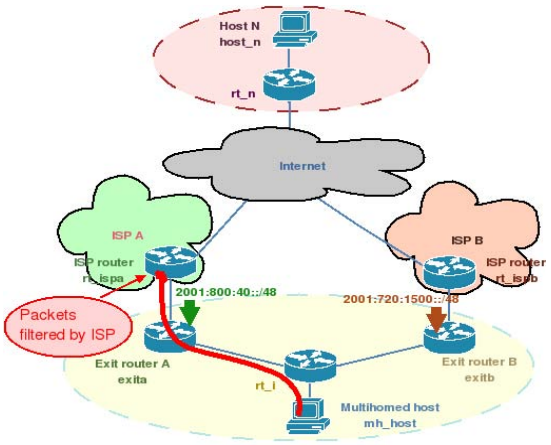
Abstract. The SAM project, aims at delivering advanced mobile services over the new generation of Internet. This paper summarizes the results regarding the provision of solutions for multihoming scenarios with mobility services in closed connection with the standard based on IPv4 and the new generation Internet based on IPv6 protocol. Moreover, the paper describes the deployment of multi-party collaboration services over mobile heterogeneous environment. Finally, the SAM project work has been focus in a real service deployment for users inside the ETSIT faculty in connection with the standard wired network.

1. Introducción

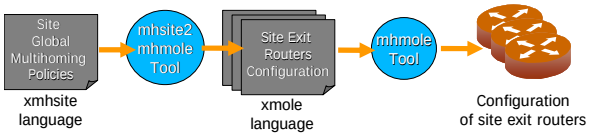
El proyecto SAM tiene como objetivo principal la provisión de servicios avanzados con movilidad y el despliegue de esos servicios de nueva generación en un entorno de uso real. El protocolo IPv6 se presenta como el catalizador que sirve para articular los servicios de la nueva generación de Internet, especialmente en aquellos aspectos relacionados con movilidad y colaboración multimedia y multiusuario.

2. Entornos con múltiples proveedores

Muchas organizaciones disponen de más de un proveedor o pertenecen a más de una organización o agrupación con accesos a la red independientes. Esto permite por un lado aumentar la tolerancia a fallos de los equipos finales y por otro acceder de forma selectiva a diferentes entornos de uso.



Lamentablemente la nueva generación de Internet basada en IPv6 ha diseñado un esquema jerárquico de direccionamiento, que es muy eficaz para tratar el enorme tamaño de la red, pero que plantea problemas en la gestión de redes con más de un proveedor de Internet. Si un mensaje se envía a un proveedor con una dirección origen de otro la respuesta puede no llegar o puede hacerlo por un camino equivocado.
El IETF ha diseñado varias soluciones; unas para ser aplicadas en cada nodo de la red y otras para filtrar en los routers de salida. En el proyecto se han estudiado e implementado algunas de las soluciones propuestas. En concreto una solución basada en nodo y compatible con los mecanismos de filtro de ingreso propuestos por IETF y basados en encaminar los mensajes que se tratan de enviar por el camino incorrecto a través de un túnel basado en direcciones *anycast.*



Así mismo se ha introducido una solución alternativa basada en configuración de políticas. La definición de las políticas se ha realizado con ayuda de un lenguaje definido en XML, que es fácil de procesar para construir configuraciones concretas de los routers y fácil de transformar para facilitar la combinación de varias políticas básicas en otras más complejas. Por todo ello la implementación que se ha mostrado sencilla de usar y fácil de desplegar.

3. Colaboración en entornos de red heterogéneos

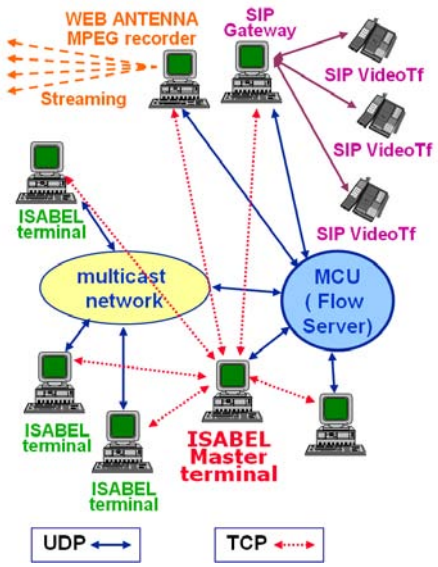
Los usuarios de Internet cada vez acceden a más servicios y aplicaciones en red utilizando un conjunto heterogéneo de terminales y redes de acceso: un ordenador portátil con conexión inalámbrica en la oficina, un móvil o PDA GPRS en la calle, un PC de sobremesa con conexión ADSL en el hogar, etc. Por tanto, surge la necesidad de proveer aplicaciones y

[1] Este trabajo ha sido parcialmente financiado por el proyecto SAM TIC2002-04531-C04-04

servicios adaptados tanto en interfaz como en funcionalidad a la variedad de mecanismos y terminales de acceso existentes.

La orientación a los Servicios Globales pretende solventar esta problemática ofreciendo servicios altamente adaptativos, que resulten accesibles independientemente de la ubicación, equipo y características de la red de acceso que esté empleando el usuario, y permitiendo incluso que dicho contexto de uso cambie dinámicamente en el transcurso de la utilización del servicio.

Uno de los campos en los que los Servicios Globales presentan un mayor impacto es el de los Servicios de Colaboración, ya que habitualmente estos presentan unos fuertes y rígidos requisitos en lo que a equipamiento y consumo de ancho de banda se refiere. Sin embargo, se trata a su vez de uno de los ámbitos de aplicación más interesantes, ya que las ventajas de aportar flexibilidad y movilidad a este tipo de aplicaciones resultan evidentes.



El proyecto SAM ha definido y validado escenarios de colaboración heterogéneos utilizando la aplicación de colaboración ISABEL, que utiliza un concepto de servicio innovador y permite interconectar audiencias distribuidas en aulas, conferencias o salas de reunión, tanto por cable como a través de redes inalámbricas.

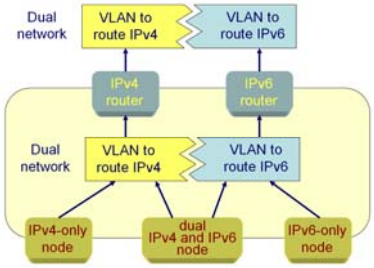
En el proyecto SAM se ha estudiado el concepto de Servicio Global implementando mecanismos para ayudar a configurar de forma transparente escenarios de colaboración sofisticados. Esto se ha conseguido con ayuda del Gestor de Configuraciones XLIM basado en un lenguaje XML, que recoge información de cada nodo participante y compone junto con la política de gestión de la conferencia que determina los diferentes modos de interacción, el entorno completo de colaboración.

ISABEL es un entorno de colaboración síncrono y el proyecto ha desarrollado algoritmos compatibles con las normas del IETF que mejoran la de calidad de servicio. Están basados en el uso de códigos FEC que permiten recuperar paquetes perdidos y son de especial importancia para entornos móviles donde es posible que se produzcan pérdidas durante los desvanecimientos existentes en enlaces inalámbricos.

4. Despliegue de servicios móviles

Aunque todavía no hay una gran proporción de servicios móviles, cada vez es más fácil disponer de dispositivos móviles con capacidad multimedia desde ordenadores portátiles a PDAs o teléfonos móviles, con capacidad de conectarse entre sí y con otros equipos fijos.

Una tarea importante del proyecto ha sido el despliegue de una red móvil en toda la Escuela de Telecomunicación que permite desplegar servicios en un entorno de transición gradual. Se ha definido un escenario de transición a IPv6 donde se ha mantenido la arquitectura de red estándar y, al mismo tiempo, se ha desplegado la nueva red IPv6.



La nueva red tiene cobertura inalámbrica en toda la escuela y permite migrar cada nodo de la red de forma independiente, sin que el resto de los equipso se vean afectados. Entre las ventajas de esta solución cabe destacar:

- La autoconfiguración que simplifica la gestión de la red.
- La posibilidad de usar IPsec para tratar los problemas de seguridad, sobre todo en el entorno móvil.
- La eficiencia en el uso de la red eliminando el uso de *broadcasting*.

Referencias

[1] Antonio Tapiador, Tomás de Miguel Maria J. Perea, Omar Walid, Marco Hernández, David Fernandez. “A simple host centric solution to multihoming Environment”. 10th Eunice Summer School and IFIP WG 6.3 Workshop (EUNICE 2004), Tampere, Finland June 14 - 16, 2004.

[2] J. Quemada, T. de Miguel, S. Pavon, G. Huecas, T. Robles, J. Salvachua, M. J. Perea, E. Moro, D. A. Acosta, J. A. Fernandez, F. Escribano, A. Diaz, J. L. Fernandez, “Isabel: An Application for real time Audience Interconnection over the Internet". Terena Networking Conference, Rodhes, June 2004.

[3] T. de Miguel, D. Fernández, J. Quemada, E. Castro. “La transición a IPv6, el nuevo protocolo de Internet”. Mundo Internet 2004, IX Congreso Nacional de Internet, Telecomunicaciones y Movilidad. Madrid, 11-13 de febrero de 2004

[4] F. Galán, D. Fernández. “VNUML: una herramienta de virtualización de redes basada en software libre. Open Source Int. Conf., Málaga 18-20 de febrero de 2004.

Servicios Avanzados con Movilidad: Provisión de Calidad de Servicio y Evaluación de los Servicios de Red

Jordi Domingo-Pascual
Centro de Comunicaciones Avanzadas de Banda Ancha (CCABA)
Departamento de Arquitectura de Comutadores
Universidad Politécnica de Catalunya
Jordi Girona 1-3. Campus Nord. Módulo D6. 08034 Barcelona
Teléfono: 934 016 981 Fax: 934 017 055
E-mail: jordi.domingo@ac.upc.edu

Abstract. *This paper presents a summary of the activities carried out within the research project SAM (Advanced Services with IP Mobility) funded by the Spanish Ministry of Science and Technology. A testbed using WLAN and Mobile IP has been set up in our laboratory at CCABA/UPC and in the laboratories of the other partners. Applications and services are tested in this distributed scenario supporting IP mobility. Most of the effort at CCABA/UPC has been dedicated to the development of measurement and monitoring tools to assess end-to-end QoS (including handovers for mobile nodes) and to the development, testing and evaluation of Fast Handovers for Mobile IPv6. Both, the measurement tools and the implementation of the Fast Handovers according to the IETF draft, are open source and are available in our web site.*

1 Introducción

Este artículo presenta un resumen de las actividades realizadas en el proyecto de investigación SAM (Servicios Avanzados con Movilidad) financiado por el Ministerio de Ciencia y Tecnología. Se trata de un proyecto coordinado pero este artículo se centra en el trabajo realizado en el Centro de Comunicaciones de Banda Ancha de la Universidad Politécnica de Cataluña (CCABA/UPC) [1]. En la página web del proyecto se publica una información más detallada del mismo [2].

Los resultados más destacados del proyecto son: la puesta en marcha de una plataforma de red experimental con movilidad IP, el desarrollo de la metodología y de las herramientas de medida de la calidad de servicio a nivel IP, y el desarrollo y validación de una implementación del Fast Handover para IPv6 de acuerdo con los borradores de RFC definidos por el IETF.

2 Plataforma de red con soporte de movilidad IP

En primer lugar se ha realizado un estudio de los protocolos y de las propuestas existentes en los grupos de trabajo del IETF sobre movilidad IP. Se han analizado las propuestas de micro-movilidad y de macro-movilidad. El resultado de este estudio previo se recoge en la publicación conjunta en un número monográfico dedicado a movilidad [3].

Se ha puesto en marcha una plataforma con soporte de movilidad IP con los objetivos de desarrollar nuevos protocolos y realizar medidas de calidad de servicio, tanto activas como pasivas.

La mayor dificultad ha sido poder encontrar tarjetas de red WLAN que se puedan utilizar como Access Point en máquinas Linux. La plataforma está basada en máquinas Linux (como routers y access points), tarjetas WLAN con el chipset de Atheros, y routers comerciales. La sincronización (necesaria para las medidas de retardo extremo a extremo) se consigue con un receptor GPS y el protocolo NTP.

3 Medidas de la calidad de servicio

Las medidas de calidad de servicio se han centrado en la latencia asociada al handover. Se ha desarrollado la herramienta PHM la cual permite capturar los mensajes de señalización y que permite determinar la duración del handover.

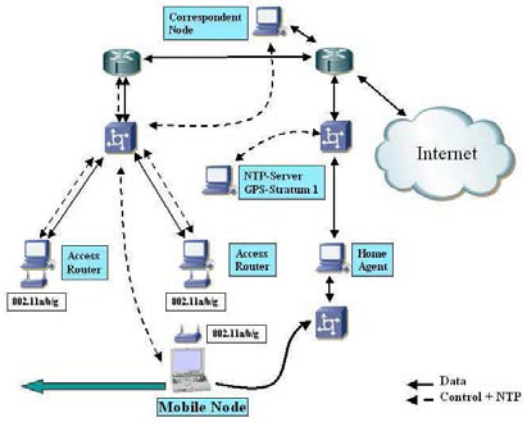


Figura 1. Plataforma de red con soporte Movilidad IP

Esta herramienta tiene soporte para IEEE 802.11, MIPv4, MIPv6 y Fast Handovers para IPv6 (FHMIP6). Adicionalmente se completa el desarrollo de la herramienta NetMeter para medir el OWD, el IPDV, y las pérdidas de paquetes antes y después del handover. Las herramientas Oreneta y NetMeter, desarrolladas en el proyecto están disponibles en código abierto y son utilizadas en otros proyectos [4, 5].

4 "Fast Handover" para IPv6

Se ha implementado el protocolo "Fast Handovers for Mobile IPv6" según los documentos de trabajo del IETF. El objetivo es reducir el tiempo de desconexión asociado al handover. El protocolo se ha implementado en C dentro del núcleo de Linux sobre una implementación existente de MIPv6. Se ha conseguido obtener los objetivos básicos del protocolo: la tasa de paqutes perdidos es cero y la latencia global del handover se reduce al hanover de nivel 2.

Las medidas realizadas y la metodología empleada, los resultados obtenidos junto que la validación de la implementación realizada se detalla en [6].

Para poder realizar un estudio de la escalabilidad de la implementación se ha utilizado el User-Mode-Linux (UML) que permite emular varios sistemas Linux en una sola máquina. El objetivo es realizar experimentos con más nodos móviles, routers de acceso, y puntos de acceso.

Agradecimientos

Este proyecto ha sido financiado por la Dirección General de Investigación del Ministerio de Ciencia y Tecnología con el contrato TIC2002-04531-C04-02. El trabajo de investigación se ha realizado conjuntamente y de forma coordinada con los grupos de investigación de la Universidad Politécnica de Madrid (UPM), la Universidad Carlos III de Madrid (UC3M) y la Universidad de Murcia (UMU). Además, se han realizador trabajos en colaboración con la Universidad Politécnica de Valencia (UPV) cuyo grupo ha participado también en las reuniones del proyecto.

Para finalizar, una mención especial para los Entes Promotores y Observadores (EPO) del proyecto, la empresa TECSIDEL y el Centro Tecnológico de Telecomunicaciones de Cataluña (CTTC), por su participación y colaboración. Cabe destacar la dedicación de un ingeniero del CTTC al proyecto SAM participando activamente en el desarrollo de los prototipos y en las reuniones de trabajo.

Referencias

[1] Centro de Comunicaciones Avanzadas de Banda Ancha de la Universidad Politécnica de Cataluña (CCABA/UPC). www.ccaba.upc.edu.

[2] Web site del Proyecto SAM. sam.ccaba.upc.es/

[3] "Movilidad IP: macromovilidad, micromovilidad, calidad de servicio y seguridad". Josep Mangues-Bafalluy, Albert Cabellos-Aparicio, René Serral Gracià, Jordi Domingo-Pascual, Antonio Gómez Skarmeta, Tomás P. de Miguel, Marcelo Bagnulo, Alberto García Martínez. NOVATICA/UPGRADE n. 167, pp. 28-32. Enero-Febrero 2004. ISSN 0211-2124 / ISSN 1684-5285.

[4] Analizador en tiempo real de la calidad de servicio en redes IP. Abel Navarro, Jordi Domingo-Pascual. Jornadas Técnicas de RedIRIS. Palma de Mallorca, 4-5/11/2003.

[5] "Active measurement tool for the EuQoS project". René Serral-Gracià, Albert Cabellos-Aparicio, Hector Julian-Bertomeu Jordi Domingo-Pascual. 3rd International Workshop on Internet Performance, Simulation, Monitoring and Measurement IPS-MoMe 2005, Warsaw, Poland.

[6] "Measurement Based Analysis of the Handover in a WLAN MIPv6 Scenario". Albert Cabellos-Aparicio1, René Serral-Gracià,Loránd Jakab, and Jordi Domingo-Pascual. 6th Passive and Active Monitoring Workshop. March 31 – April 1, 2005, Boston MA (USA). C. Dovrolis (Ed.): PAM 2005, LNCS 3431, pp. 207–218, 2005.

Servicios Avanzados con Movilidad (TIC2002-04531-C04-03) – Universidad Carlos III de Madrid

Alberto García-Martínez, Marcelo Bagnulo
Departamento de Ingeniería Telemática. Universidad Carlos III de Madrid
Avda. Universidad s/n. 28911 – Leganés (Madrid)
E-mail: alberto@it.uc3m.es, marcelo@it.uc3m.es

Abstract. *En esta nota se recogen las contribuciones realizadas en el marco del proyecto Servicios Avanzados con Movilidad (TIC2002-04531-C04-03). El desarrollo del proyecto se ha centrado en cubrir dos requisitos identificados para las redes IPv6: la definición de un modelo de implantación escalable de multihoming, y la apropiada provisión de servicios de nomadismo mediante el uso de DHCPv6. Además se ha desarrollado una maqueta de comunicaciones multimedia basada en movilidad sobre IPv6.*

1 Introducción

IPv6 es uno de los protocolos alrededor de los que se espera que se articulen servicios de Nueva Generación, servicios entre los que destacan aquellos relacionados con movilidad. Este proyecto se ha centrado en dos temas de interés para IPv6: el soporte de múltiples proveedores, y el soporte de nomadismo mediante la implementación de DHCPv6. En las secciones siguientes detallamos los problemas y las contribuciones realizadas por el grupo de trabajo del proyecto SAM en la U. Carlos III de Madrid.

2 Soporte de Múltiples Proveedores para IPv6

La provisión de soporte de múltiples proveedores (*multihoming*) permite que una red o un equipo final se puedan beneficiar de mejor tolerancia a fallos o de nuevas capacidades de ingeniería de tráfico al disponer de múltiples proveedores hacia Internet. Un soporte de este tipo es demandado por aplicaciones que ofrecen servicios de comunicaciones multimedia y de colaboración, como por ejemplo por ISABEL que se está siendo utilizada como aplicación de referencia por el resto de los miembros del proyecto coordinado. La solución utilizada actualmente en IPv4 se basa en el uso del protocolo BGP de encaminamiento interdominio. No obstante, esta solución presenta importantes limitaciones que no hacen deseable su implantación en IPv6 tal y como se utiliza ahora, ya que incrementa la carga en el sistema de rutas, e impide que redes pequeñas – incluyendo usuarios residenciales – puedan beneficiarse del multihoming.

Si bien se venía reconociendo desde hace tiempo la necesidad de encontrar una solución para multihoming en IPv6 con mejores características que la actual de IPv4, sólo se habían encontrado soluciones parciales para casos particulares tales como la presentada en RFC 3178, que ha sido automatizada por los autores en [1]. La arquitectura de una solución más completa no se ha definido hasta fechas muy recientes, debido principalmente a la necesidad de ofrecer protección frente a posibles nuevos problemas de seguridad.

Para el soporte de multihoming en IPv6 proponemos dos conjuntos de herramientas [2] que se basan en la posibilidad que habilita IPv6 de disponer de una dirección por cada proveedor (configuración conocida como *multidireccionamiento*).

El primer conjunto de herramientas facilita que se puedan establecer nuevas comunicaciones en caso de fallo mediante un esquema de encaminamiento intradominio en el que se tenga en cuenta también la dirección fuente para el reenvío de paquetes y mediante una adaptación del mecanismo de Selección de Direcciones por Defecto (RFC 3484).

El segundo conjunto de herramientas permite preservar comunicaciones ya establecidas en caso de fallo. Este problema, más complejo que el anterior, se aborda en una nueva subcapa a nivel de red [3] que gestiona dinámicamente, mediante un protocolo específico, los múltiples localizadores por proveedor disponibles en el esquema de multidireccionamiento. Las herramientas de las que dispondría esta capa [4] permitirían la gestión de los localizadores, la detección de fallos en caminos dados, encontrar caminos alternativos y poder forzar su uso, e identificar un flujo independientemente del camino utilizado. Para evitar la aparición de nuevos ataques de seguridad que permitieran suplantar de forma fácil a un usuario o realizar ataques de denegación de servicio, se ha propuesto el uso de Direcciones Basadas en Hash (*Hash Based Addresses*, HBA [5], [6]), que logran este objetivo estableciendo la ligazón de un conjunto de localizadores con una identidad dada sin requerir operaciones de clave asimétrica con alto coste computacional. El grupo de trabajo del proyecto SAM de la Universidad Carlos III ha contribuido muy activamente a la solución que se está discutiendo actualmente en el IETF en los grupos de trabajo `multi6` y `shim6`.

3 Soporte de Nomadismo mediante DHCPv6

Otro déficit reconocido a IPv6 ha sido la lentitud en la estandarización de DHCPv6. Si bien IPv6 ha ofrecido desde sus primeros estándares mecanismos de autoconfiguración mejorados respecto a IPv4, estos mecanismos no siempre son suficientes si se requiere configuración de parámetros de niveles superiores al nivel de red (como identificación de los servidores DNS disponibles), o si un servicio requiere algún tipo de configuración más complejo que el actualmente definido. Si definimos *nomadismo* como la posibilidad de obtener conectividad en una red para la que un equipo no está especialmente configurado, distinguiéndolo de *movilidad* en que en nomadismo no es un requisito el mantenimiento de las comunicaciones previamente establecidas en otra red, coincidiremos en que DHCPv6 es requerido para dar un soporte genérico al nomadismo.

En la U. Carlos III de Madrid se ha desarrollado un cliente, servidor y repetidor (*relay*) de DHCPv6 para Linux y FreeBSD conforme a los estándares RFC 3315, RFC 3736, RFC 3646, siendo el software fácilmente extensible. Para facilitar el desarrollo de configuraciones en el servidor se ha desarrollado adicionalmente un interfaz gráfico.

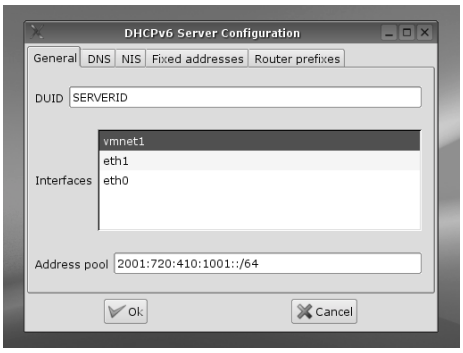


Figura 1. Captura de pantalla de configuración del servidor.

4 Conclusiones

La principal contribución realizada en el marco de este proyecto se centra en el soporte de múltiples proveedores para entornos IPv6 mediante HBAs, solución actualmente desarrollada en los correspondientes grupos de trabajo del IETF.

Por otro lado, se ha desarrollado un cliente/servidor de DHCPv6 que ofrece facilidades de configuración a clientes nómadas. Además de la funcionalidad operativa, el disponer de un software de DHCPv6 fácilmente extensible permite la realización de experiencias de configuración de otros servicios tales como la tabla de políticas de selección de direcciones definidas en RFC 3484.

En relación con el protocolo MIPv6, se han realizado dos contribuciones: La primera, analizar el posible uso de MIPv6 como herramienta de soporte para multihoming, llegando a la conclusión de que no parece apropiado, a pesar de que la comunidad científica había destacado similitudes entre los problemas de movilidad y multihoming [7]. La segunda es una mejora en la especificación del protocolo de movilidad en IPv6 para permitir preservar comunicaciones previamente establecidas cuando el camino entre el nodo móvil (MN) y el nodo correspondiente (CN) a través de la *Home Address* (HoA) no se encuentra disponible [8].

Finalmente, se ha colaborado con otros miembros del proyecto coordinado en la realización de experiencias de teleeducación y de movilidad basadas en MobileIPv6.

Agradecimientos

Agradecemos a Carlos Izquierdo su contribución al desarrollo del software de DHCPv6.

Referencias

[1] M. Bagnulo, J. F. Rodríguez-Hervella, A. García-Martínez, A. Azcorra. Multi-Homing Tunnel Broker. IEEE Euromicro 2004, Rennes, Francia. Págs 282-289. Septiembre 2004.

[2] M. Bagnulo, A. García-Martínez, A. Azcorra. Herramientas para la provisión de multihoming en IPv6. Novática. Abril 2005. También publicado como “Tools for IPv6 Multihoming”, Revista Upgrade.

[3] E. Nordmark, M. Bagnulo. Multihoming L3 Shim Approach. draft-ietf-multi6-l3shim-00. Enero 2005.

[4] M. Bagnulo, J. Arkko. Functional decomposition of the M6 protocol. draft-ietf-multi6-functional-dec-00. Diciembre 2004

[5] M. Bagnulo. Hash Based Addresses (HBA). draft-ietf-multi6-hba-00. Diciembre 2004.

[6] M. Bagnulo, A. García-Martínez, A. Azcorra. Efficient Security for IPv6 Multihoming. Pendiente de publicación en ACM Computer Communications Review.

[7] M. Bagnulo, A. García-Martínez, I. Soto. Application of the MIPv6 protocol to the multi-homing problem. draft-bagnulo-multi6-mnm-00. Julio 2003.

[8] M. Bagnulo, A. García-Martínez, I. Soto. Preserving MIPv6 communications when the HoA becomes unreachable. draft-bagnulo-mobileip-unreachable-hoa-00. Junio 2003.

Redes ad hoc, Comunicaciones Multimedia y Control de Acceso en el Marco del Proyecto SAM[1]

Antonio F. Gómez Skarmeta, Pedro M. Ruiz Martínez
Depto. Ingeniería de la Información y las Comunicaciones. Universidad de Murcia
Facultad de Informátia. Campus de Espinardo s/n.
30071 – Espinardo (Murcia)
Teléfono: 968 36 46 07 Fax: 968 36 41 51
E-mail: {pedrom,skarmeta}@dif.um.es

Abstract. *The SAM project, aims at delivering advanced mobile services. This paper describes the results regarding the provision of multimedia multi-party communicaations in ad hoc networks and the results on access control based on AAA and authorization attributes. For efficient multicasting in ad hoc networks we have developed the MMARP protocol. Multimedia communications are handled through the use of adaptive multimedia applications and we also provided an enhanced policy-based access control architecture.*

1 Introducción

El proyecto SAM tiene como objetivo principal la provisión de servicios avanzados con movilidad. En este artículo describiremos nuestras propuestas para la provisión de comunicaciones multidestino eficientes en redes ad hoc, la provisión de servicios multimedia en entornos de red cambiantes y la provisión de mecanismos de control de acceso en redes móviles.

En lo referente a las comunicaciones multidestino, a lo largo del proyecto se han definido extensiones al protocolo MMARP[1] para proveerlo de una mejor interconexión a redes fijas, extensiones de seguridad y para reducir su sobrecarga de datos. Por último, hemos ofrecido una arquitectura mejorada para el control de acceso a redes móviles basado en políticas mediante el empleo de SAML y XACML[2]. En las siguientes secciones describimos todos estos trabajos y presentamos algunas conclusiones.

2 Extensiones a MMARP

El protocolo MMARP fue pionero en la provisión de comunicaciones multicast entre redes ad hoc y redes fijas. Tal cual muestra la Fig.1, MMARP define el uso del protocolo IGMP (o MLD en IPv6) para ofrecer una integración con redes fijas. Además, MMARP define su propio mecanismo para el descubrimiento de gateways hacia la red fija, así como el uso de MIGs, para realizar las funciones de interacción. Sin embargo, actualmente se han propuesto en el IETF diferentes mecanismos de interconexión entre redes ad hoc y redes fijas. Dado que en general es deseable usar el mismo mecanismo de interconexión para los protocolos unicast y los protocolos multicast, en este proyecto hemos

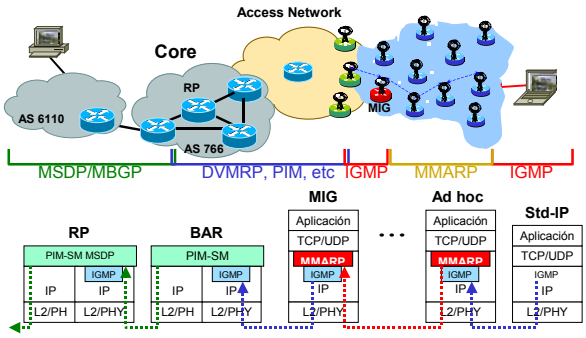


Figura. 1. Arquitectura multicast para extensiones ad hoc

integrado la propuesta de Jelger[3], con la implementación de MMARP realizada por nuestro grupo. Por supuesto, la implementación ha sido extendida para funcionar en redes IPv6. De este modo, los nodos MMARP descubren rutas para encaminar tráfico multicast hacia Internet del mismo modo que se descubren las rutas unicast.

Otra de las propuestas de mejoras introducidas sobre la definición inicial de MMARP ha sido la de autenticar los mensajes de control. Para ello, usando criptografía asimétrica se propone ofrecer integridad de los mensajes, mientras que la autenticación se consigue mediante CGA. Para ello, un nodo MMARP configurará su prefijo de red a partir del protocolo de Jelger y su identificador del interfaz lo generará criptográficamente con CGA.

Por último, se ha optimizado la construcción de la malla de encaminamiento multicast empleada por MMARP. En concreto, empleamos un enfoque adaptable por el que la fiabilidad de la malla se adapta a la movilidad de la red. Nuestros resultados[4] muestran que el esquema mejorado es capaz de reducir la sobrecarga de datos en un 40% y la latencia en un 10%.

[1] Este trabajo ha sido parcialmente financiado por el proyecto SAM TIC2002-04531-C04-04

3 Control de Acceso

En escenarios móviles, cada usuario final pertenece a un dominio origen, donde posee un conjunto de atributos según el papel que juega. En el proyecto SAM se ha realizado una aproximación al control de acceso a la red basada en certificados de identidad X.509 y en atributos de autorización. Esta aproximación presenta algunos de los retos derivados de la integración de sistemas de autorización existentes, y un sistema de autorización flexible, escalable y manejable. Nuestra propuesta está basada en los estándares SAML y XACML [2], que serán usados para expresar políticas de control de acceso basadas en atributos, sentencias y protocolos de autorizaciónCuando el usuario móvil solicita acceso a la red en otro dominio diferente al suyo, la solicitud es recogida por el servidor AAA del dominio visitado, el cual genera una consulta para obtener los atributos asociados al usuario desde la autoridad responsable de gestionarlos. Finalmente, una vez recogidos estos atributos, el servidor AAA envía una consulta a un PDP(Policy Decision Point), el cual responde indicando si los atributos del usuario satisfacen la política de control de acceso del dominio. Además, esta política puede establecer el conjunto de obligaciones derivadas de esta decisión, por ejemplo atributos relacionados con QoS, opciones de seguridad, etc. Como se verá, este esquema general puede implantarse en un escenario mono o multi-dominio. La Fig. 2 muestra los principales componentes de este escenario.

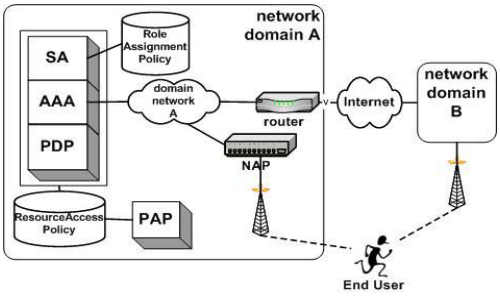


Figura 2. Arquitectura de control de acceso propuesta

En nuestra arquitectura la identidad de los usuarios finales se gestiona mediante certificados X.509 emitidos por la autoridad de certificación (CA) del dominio origen. Además, incluirá atributos adicionales empleados para determinar sus derechos de acceso. El servidor AAA usará el protocolo DIAMETER para comunicarse con otros servidores AAA, así como para negociar la generación de sentencias SAML con el SA (Source Authority) y con los PDPs para generar decisiones de autorización. Cada SA tendrá una política de asignación de roles basada en XACML. Además, se definen dos modelos de operación. En el modelo ¨push¨, el usuario contactará con la SA para obtener sus roles, antes de solicitar el acceso a la red. En el podelo ¨pull¨es el servidor AAA local o remoto el que contacta directamente a la SA.

Para poder soportar escenarios multidominio en modo ¨pull¨, hemos tenido que extender el protocolo DIAMETER para permitir que los servidores AAA puedan enviar solicitudes y respuestas que contengan información de SAML. Esta extensión que denominamos DIAMETER-SAML define nuevos atributos para llevar carga útil SAML. Este esquema completo en modo pull se muestra en la Fig. 3.

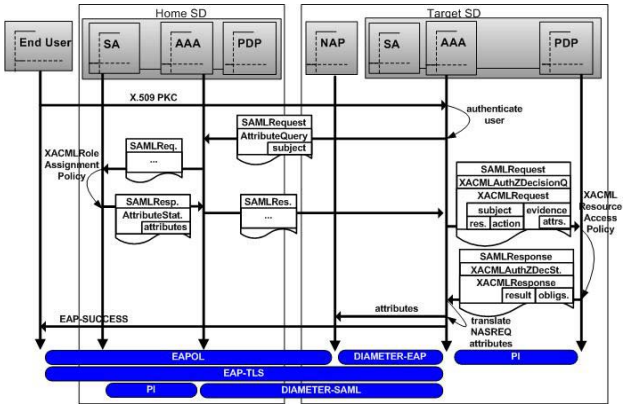


Figura 3. Modelo pull

Como se aprecia, una vez el servidor AAA ha validado la identidad del usuario, chequea si el usuario tiene permiso para acceder a la red. Para ello, el servidor AAA obtiene los roles del usuario de la SA, que al recibir la solicitud obtiene los roles usando la política de asignación de roles y los envía en una respuesta firmada al servidor AAA. Para esa comunicación multidominio se emplea DIAMETER-SAML. Ahora el servidor AAA consulta al PDP para obtener una decisión de autorización. Esta consulta incluye el recurso solicitado, el usuario que hace la solicitud y la acción a realizar sobre el recurso. Finalmente, el servidor AAA al recibir la respuesta del PDP habilita la conexión de red conforme a las obligaciones indicadas por el PDP en la respuesta.

Referencias

[1] Pedro M. Ruiz, Antonio F. Gómez-Skarmeta, “Integrated IP Multicast in Mobile Ad Hoc Networkis with Multiple Attachments to Wired IP Networks”, 15th IEEE Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC-2004), Barcelona, 5-8 Sept. 2004.

[2] S., Godik and T. Moses. “OASIS eXtensible Access Control Markup Language (XACML) Versión 1.0, February 2003. OASIS Standard.

[3] C. Jelger, T. Noel, A. Frey IETF draft-jelger-MANET-gateway-autoconf-v6-01.txt, Oct.2003.

[4] Pedro M. Ruiz, Juan A. Botía, Antonio Gómez-Skarmeta, "User-aware Adaptive Applications for Enhanced Multimedia Quality in Heterogeneous Networking Environments", Lecture Notes in Computer Science, vol. 3079, June 2004.

Índice onomástico de autores

Se arrebola el sol.
Donde muere la vista
pervive el verso.

Este libro se terminó de imprimir en los talleres de Tórculo Artes Gráficas
de Santiago de Compostela, el 20 de julio de 2005.