

# Diseño e Implementación de un Servicio de TV/Cachecasting a móviles sobre DVB-H

*Autor:* Francisco Fraile Gil

*Director:* Juan Carlos Guerri Cebollada

## *Resumen*

En este trabajo se presenta la arquitectura de un servicio de descarga de programas de televisión (*TV/Cachecasting*) a terminales móviles sobre redes IP multicast DVB-H. Los resultados obtenidos muestran el tiempo medio de acceso a los archivos para varias disciplinas de envío de ficheros en el servidor y las prestaciones de algoritmos de gestión de caché que utilizan un modelo de las preferencias de los usuarios para implementar políticas de reemplazo en memoria. Además, se proporcionan guías de implementación para el despliegue de este tipo de servicios.

## *Abstract*

This work presents the architecture of a TV program download service to mobile terminals (*TV/Cachecasting*) over IP multicast DVB-H Networks. Results show the overall access time to programs for different server policies and the performance of cache management algorithms that use models of the user preferences to implement cache replacement policies. Furthermore, this work provides implementation guidelines for the deployment of this kind of services.

Autor: Fraile Gil, Francisco, e-mail: [ffraile@iteam.upv.es](mailto:ffraile@iteam.upv.es)

Director: Guerri Cebollada, Juan Carlos, e-mail: [jcguerri@dcom.upv.es](mailto:jcguerri@dcom.upv.es)

Fecha de Entrega: 26-03-2009

**ÍNDICE**

|  |           |
|--|-----------|
| <b>I. INTRODUCCIÓN</b> .....   | <b>3</b>  |
| I.1. MOTIVACIÓN.....   | 3         |
| I.2. OBJETIVOS .....   | 4         |
| I.3. CONTRIBUCIONES.....   | 5         |
| I.4. ESTADO DEL ARTE DE LOS SISTEMAS DE DISTRIBUCIÓN DE CONTENIDO<br>MULTIMEDIA A MÓVILES..... | 6         |
| <b>II. MULTIPLEXACIÓN DE OBJETOS EN SISTEMAS DE DIFUSIÓN</b> .....                             | <b>9</b>  |
| <b>III. GESTIÓN DE CACHÉ Y PERSONALIZACIÓN</b> .....   | <b>12</b> |
| <b>IV. ESTUDIO DE LA ARQUITECTURA DEL SERVICIO</b> .....                                       | <b>13</b> |
| IV.1. EL PROTOCOLO FLUTE.....  | 14        |
| IV.2. LA GUÍA DE SERVICIO OMA BCAST SERVICE GUIDE.....   | 18        |
| IV.1. IMPLEMENTACIÓN DE ALGORITMOS DE REEMPLAZAMIENTO EN CACHÉ<br>BASADOS EN PIX.....          | 20        |
| <i>IV.1.1. IMPLEMENTACIÓN A PARTIR DE UN RECOMENDADOR HEURÍSTICO ...</i>                       | <i>23</i> |
| <i>IV.1.2. IMPLEMENTACIÓN A PARTIR DE UN RECOMENDADOR BAYESIANO .....</i>                      | <i>24</i> |
| <b>V. MÉTODOS Y RESULTADOS</b> .....   | <b>24</b> |
| V.1. EVALUACIÓN ANÁLITICA DE LA MULTIPLEXACIÓN DE OBJETOS .....                                | 24        |
| V.2. EVALUACIÓN DE LAS PRESTACIONES DEL ALGORITMO VIRTUAL CLOCK<br>MODIFICADO .....            | 28        |
| V.3. EVALUACIÓN DE LOS MECANISMOS DE CACHING.....  | 29        |
| V.4. RESULTADOS DE LAS MEDIDAS EN LABORATORIO .....  | 30        |
| V.5. ANÁLISIS DE LOS RESULTADOS.....   | 31        |
| <b>VI. CONCLUSIONES Y TRABAJO FUTURO</b> .....   | <b>33</b> |
| <b>AGRADECIMIENTOS</b> .....   | <b>35</b> |
| <b>REFERENCIAS</b> .....   | <b>36</b> |
| <b>ANEXOS</b> .....  | <b>38</b> |

## I. INTRODUCCIÓN

### I.1.MOTIVACIÓN

A la vez que comienza el despliegue a nivel global de sistemas de difusión de contenido multimedia a móviles a través de tecnologías como DVB-H, T-DMB, MediaFlo o MBMS, la industria audiovisual se plantea cómo deben ser adaptados los servicios televisivos al entorno móvil. En la actualidad, está generalmente aceptado que la producción de contenidos debe cuidarse para proporcionar una experiencia satisfactoria durante la proyección en terminales móviles, dado que los planos largos, los movimientos rápidos de cámara o las escenas sobrecargadas no se visualizan de manera satisfactoria en pantallas de pequeño tamaño. Además, las pruebas comerciales de los servicios de TV móvil han demostrado que los intervalos de uso son de pocos minutos de duración media, mucho más cortos que en la televisión tradicional. Es lógico pensar que la duración de los programas de TV a móviles debe adaptarse en concordancia con los patrones de uso y por tanto, la duración de los mismos no ha de exceder de unos pocos minutos.

Las pruebas piloto también ponen de manifiesto que si los tiempos de uso disminuyen, las expectativas en los contenidos aumentan y los usuarios esperan encontrar contenidos adaptados a sus preferencias cuando acceden al servicio. Además, el uso de la TV móvil es mucho más errático que la TV tradicional, y resulta complicado encontrar franjas *prime time* alrededor de las cuales diseñar la programación. Estos hechos corroboran que la edición de la programación de un servicio de TV a móviles es una tarea complicada.

A la hora de decidir qué contenidos se deben transmitir y cuándo, resulta mucho más complicado satisfacer las expectativas de los usuarios en movilidad si la transmisión de los programas se realiza de manera lineal, como en los sistemas tradicionales de difusión de contenido en tiempo real. Sin embargo, los métodos de transporte especificados para los sistemas de difusión a móviles basados en tecnología IP, definidos en estándares como OMA BCAST o DVB IPDC, ofrecen otra alternativa. Estos estándares proporcionan dos métodos para la distribución de contenidos sobre los que construir los servicios: por un lado, el protocolo *Real Transport Protocol*, (RTP), que se utiliza para la transmisión de contenidos en tiempo real; y por otro, el protocolo *File Delivery over Unidirectional Transport*, FLUTE, para la distribución de archivos en diferido. Cabe señalar que, en este contexto, transmisión en diferido significa que los usuarios no visualizan el programa en el intervalo de tiempo en que su emisión utiliza recursos en la red de difusión (al menos no necesariamente). En lugar de reproducir los contenidos a la vez que se emiten, éstos se van almacenando en la memoria del terminal hasta que se completa la descarga. Así, los contenidos transmitidos por FLUTE se caracterizan por dos intervalos de tiempo: un intervalo en el que se habilita la transferencia en la red, denominado **ventana de transmisión**; y otro intervalo de tiempo en el que se programa su reproducción por los terminales, denominado **ventana de reproducción**.

Los servicios que utilizan el protocolo de transporte FLUTE para la transmisión de los programas de TV se denominan servicios de TV/*Cachecasting* y son el objeto de estudio de esta Tesina, porque proporcionan mecanismos para facilitar la edición de la programación de los servicios de TV a móviles, de acuerdo a las expectativas de los usuarios.

Los servicios propuestos en esta Tesina hacen uso de los recursos de red para la emisión de los programas mediante sesiones de transferencia de ficheros FLUTE. Esto permite utilizar la capacidad total asignada al servicio para ofrecer a los usuarios contenidos en una programación adaptada a sus preferencias, en contraste con los mecanismos de transmisión tradicionales, en los que la programación es lineal con el tiempo e igual para todos los usuarios.

Dicho de otro modo, al utilizar los recursos de red para transmitir los programas de una manera no lineal, el operador del servicio tiene la oportunidad de adaptar la oferta de contenidos a las preferencias de los usuarios. El operador puede integrar en la programación los contenidos que se ofrecen por FLUTE en cualquier intervalo posterior al inicio de la sesión de transferencia de ficheros, de manera que el editor de programación podría decidir que múltiples programas estén disponibles para su visualización por parte de los usuarios del servicio. Los terminales seleccionan, de entre los contenidos que se programan simultáneamente, aquellos que se ajustan mejor a las preferencias del usuario. Para realizar esta selección, la aplicación cliente utiliza las descripciones sobre los programas que se incluyen en la Guía de Servicio. *De esta manera, se ofrece al editor de programación la posibilidad de generar programaciones adaptadas a las preferencias de los usuarios, utilizando únicamente un canal de difusión unidireccional, sin necesidad de un canal de comunicación con los usuarios.*

## I.2.OBJETIVOS

El objetivo principal de esta Tesina es *evaluar las mejoras en las prestaciones de un servicio de TV/Cachecasting introducidas por la aplicación de multiplexación de objetos y por la introducción de mecanismos de gestión de caché basados en una estimación de las preferencias de los usuarios sobre los contenidos*. También se proporcionan guías de implementación para introducir estas mejoras en un servicio de TV/*Cachecasting* para terminales móviles basado en el estándar DVB-H. Para conseguir estos objetivos, se plantean los siguientes hitos:

- i. **Estudio de la arquitectura del Servicio TV/*Cachecasting* sobre DVB-H.** Como primer objetivo, se definirá la arquitectura del servicio, mostrando los parámetros de configuración de los principales componentes involucrados (nivel de enlace DVB-H, FLUTE, ESG, etc.).
- ii. **Análisis de la viabilidad del Servicio.** A partir de un estudio analítico del rendimiento del sistema y de un modelo software, se realizarán simulaciones para evaluar bajo qué condiciones se ofrecen niveles de Calidad de Servicio aceptables.

- iii. **Implementación de un prototipo del Servicio.** Utilizando los recursos del laboratorio DVB-H del Instituto de Telecomunicaciones y Aplicaciones Multimedia (iTEAM), se desarrollará un prototipo del servicio para realizar pruebas en el laboratorio.
- iv. **Pruebas de Evaluación y Optimización.** Los resultados analíticos obtenidos se contrastarán con pruebas de laboratorio, para cuantificar los márgenes de implementación (desviación de los resultados con respecto a los obtenidos en el análisis de viabilidad) que presenta el servicio sobre DVB-H.

### I.3.CONTRIBUCIONES

La primera fase de este trabajo de Tesina se destinará a realizar un análisis exhaustivo de la topología de los servicios mixtos TV/*Cachecasting* sobre DVB-H. El estudio enfatizará en los aspectos tecnológicos más relevantes para el caso de estudio, como la transferencia de ficheros sobre redes DVB-H sobre el protocolo FLUTE, el modelo de datos de la ESG para los servicios mixtos TV/*Cachecasting* y las funciones de personalización que se pueden aplicar a través de ella.

En base a este pre-estudio, se evaluará la introducción de dos mejoras en la plataforma de descarga de ficheros sobre DVB-H del Grupo de Comunicaciones Multimedia (COMM) del Instituto de Telecomunicaciones y Aplicaciones Multimedia (iTeam).

La primera se denomina *multiplexación de objetos*, que es una propuesta descrita en [1] para mejorar las prestaciones de los servicios de descarga de objetos sobre redes de difusión. En dicho estudio se obtiene un límite mínimo teórico del tiempo medio de acceso a un objeto para cierta configuración de una transmisión cíclica de objetos sobre una red de difusión. La tasa a largo plazo utilizada para la transmisión de cada objeto depende del tamaño de los objetos y de una medida empírica de su popularidad. Los primeros resultados irán orientados a evaluar dicho límite. Después se propondrá una implementación de la multiplexación de objetos derivada del algoritmo de *Fair Queuing* conocido como *Virtual Clock*, que también ha sido planteada en [1]. Se evaluarán los márgenes de implementación al utilizar *Virtual Clock* para ajustar la tasa de transmisión a largo plazo para cada objeto a través de un modelo del sistema desarrollado en *Matlab*.

La segunda mejora que se planteará será la de introducir un *mecanismo de gestión de caché en el cliente que utilice una estimación de las preferencias de los usuarios para almacenar los contenidos en la memoria del terminal*. El modelo software utilizado en el apartado anterior se modificará para cuantificar la mejora que introduce el uso de memoria caché en el tiempo medio de acceso. Después, se proporcionarán guías de implementación para el algoritmo de gestión de caché, que utilizará la información disponible en la Guía Electrónica de Servicios sobre los programas, además del historial de accesos a otros contenidos, para cuantificar el valor de cada objeto al aplicar políticas de reemplazo en caché.

Este documento se estructura de la siguiente manera: La siguiente sección incluye un pequeño resumen sobre el estado del arte en redes de difusión de contenido a móviles. Los siguientes

capítulos se dedican a presentar las mejoras que se quieren introducir en el sistema, la multiplexación de objetos y el uso de caché en redes de difusión. Después se muestra la arquitectura del servicio para el caso de uso, el protocolo FLUTE, sobre el que se construyen los servicios de descarga de ficheros y por último lugar se muestra la ESG del estándar OMA BCAST, elegido para la implementación de la gestión de caché. La siguiente sección se dedica a los métodos y resultados. Se presentarán los resultados obtenidos al evaluar el límite teórico. Después, se presentan las prestaciones del algoritmo *Virtual Clock* modificado y las mejoras introducidas al utilizar caché. Los últimos resultados consisten en guías de implementación para introducir funciones de personalización en sistemas de descarga de ficheros unidireccionales. Por último, se incluyen algunas conclusiones elaboradas a partir de los resultados.

#### I.4. ESTADO DEL ARTE DE LOS SISTEMAS DE DISTRIBUCIÓN DE CONTENIDO MULTIMEDIA A MÓVILES

La recepción de programas de televisión en movilidad es posible desde hace unas décadas, dada la disponibilidad de receptores de TV analógica portátiles. Sin embargo, la utilización de este tipo de terminales se puede considerar como anecdótica. Por esta razón, es difícil justificar el interés que existe en la actualidad en los servicios de televisión a móviles por el mero hecho de disponer de los servicios televisivos en movilidad. La verdadera motivación que ha propiciado el despliegue de estos servicios es la de disponer de la televisión en el teléfono móvil, o en cualquier otro reproductor multimedia que lleve consigo el usuario, ya que hasta recientemente, era el único servicio de telecomunicaciones que no estaba disponible en este tipo de terminales, donde la tendencia es la de integrar el mayor número de funciones y servicios posibles. *El servicio de televisión ha generado una gran expectación en la industria, puesto que se considera un buen incentivo para atraer nuevos clientes, a través de una oferta triple: TV, datos y telefonía.*

La evolución tecnológica de los dispositivos también ha favorecido el consumo de contenido multimedia en los terminales móviles. Los nuevos terminales permiten la reproducción de contenido audiovisual con una resolución aceptable y sonido estéreo de calidad, lo que motiva una demanda por parte de los usuarios de servicios de telecomunicaciones que ofrezcan estos contenidos.

Existen tres alternativas para el despliegue de servicios de distribución de contenido audiovisual a móviles: i) Utilizar las redes de transmisión de datos a móviles, ii) posibilitar la recepción de los servicios de difusión de televisión digital en los móviles o iii) emplear redes de difusión dedicadas para servicios de TV a móviles.

En la actualidad, los servicios de *streaming* de vídeo a través de redes de transmisión de datos a móviles son el mecanismo más extendido para la transmisión de contenido audiovisual a estos terminales. Estos servicios se ofrecen desde redes de datos 2.5G como GPRS y EDGE, pero también desde redes 3G y HSDPA. Sin embargo, las tasas de transmisión que ofrecen las redes 2.5

no permiten la emisión de vídeo con una Calidad de Servicio aceptable. La provisión de servicios de *streaming* de vídeo con suficiente Calidad de Servicio es sólo viable a partir de las redes 3G. Normalmente, los codecs utilizados son H.264 y AAC, el transporte del vídeo se realiza a través de RTP y el control del flujo con el protocolo RTSP. Esta arquitectura de servicio permite ofrecer contenido audiovisual bajo demanda. En la mayoría de los países europeos donde están disponibles los servicios de telefonía de tercera generación, se ofrecen programas de televisión adaptados a móviles a través de esta tecnología. También se ofrecen a través de *streaming* los servicios de televisión tradicional, generalmente canales *Premium*. En el modelo de negocio más extendido, los operadores del servicio de 3G pagan a los proveedores de contenidos televisivos por ofrecer sus programas a los usuarios de redes 3G. Los operadores de red pueden tarificar a los usuarios por la descarga de datos, aunque cada vez más, se aplican tarifas planas. En este caso los beneficios se registran por publicidad, en el portal 3G del operador y porque los contenidos constituyen un reclamo para los clientes. El principal problema que presenta este modelo es la escalabilidad, dado que la capacidad de las redes 3G es limitada y compartida con otros servicios de datos. Con muchos usuarios, el servicio no es viable por la capacidad limitada de la red y con pocos, no es viable por los elevados costes que implica la adquisición de contenidos.

Para mejorar el rendimiento de la red 3G a la hora de ofrecer información a múltiples usuarios simultáneamente, se introduce la mejora *Multimedia Broadcast Multicast Service* (MBMS), que permite la utilización de IP Multicast sobre redes UMTS y GSM. Los estándares *Mobile Broadcast Services Enabler Suite* (BCAST), desarrollados por el consorcio *Open Mobile Alliance* (OMA), especifican los protocolos de transporte para servicios de *streaming* y de transferencia de ficheros sobre MBMS. Para los servicios de televisión, BCAST sobre MBMS permite la transmisión de flujos RTP multicast a los terminales móviles de una celda sin que el número de usuarios que sintonizan el servicio afecte a la capacidad de red requerida por el servicio. BCAST especifica que los servicios de transferencia de ficheros deben utilizar el protocolo FLUTE.

Sin embargo, el número total de servicios que se pueden ofrecer sobre MBMS es limitado. La capacidad máxima para una celda es de 1.7Mbits/seg, que da lugar a un número máximo de 6 servicios de televisión a 256kbps en una frecuencia de UMTS. Se considera que en la mayoría de los mercados, una oferta de estas características es demasiado limitada y que un servicio de televisión a móviles debe incluir un número mayor de canales.

Existen diversos sistemas dedicados a la transmisión de contenido multimedia a móviles que ofrecen una capacidad superior a MBMS, basados en las normas *Digital Multimedia Broadcast* (DMB) y *Digital Video Broadcast* (DVB).

Existen dos variantes de DMB, DMB-T (Terrestre) y DMB-S (Satélite), ambas muy extendidas en Corea del Sur, país de origen de las especificaciones. Las normas DMB, basadas en la norma europea *Digital Audio Broadcast* (DAB) utilizan codificación MPEG-4 - AAC y MPEG-TS para el transporte. Existen una considerable cantidad de terminales móviles que implementan estas

normas, ya que en Corea del Sur se ofrecen servicios sobre DMB desde 2005. En Europa, existen pilotos basados en T-DMB en Alemania, Francia, Reino Unido y Noruega. Los gobiernos de estos dos últimos países han lanzado concursos para la utilización de recursos espectrales para ofrecer servicios basados en T-DMB. T-DMB utiliza canales DAB (1.5 MHz en las bandas L y VHF). En Corea del Sur T-DMB utiliza la banda VHF (por lo que los terminales integran antenas telescópicas, algo poco común en el mercado europeo), pero en Europa sólo se considera viable su implantación en la banda L, por la canalización (8MHz) y la saturación de la banda VHF en el continente. Para servicios de transferencia de datos, el sistema T-DMB adopta la norma *Multimedia Object Transfer Protocol* (MOT), definida para la transferencia de archivos sobre DAB. Las especificaciones también definen la interfaz de programación que deben ofrecer los clientes de T-DMB para el desarrollo de aplicaciones que utilicen esta tecnología. La norma define dos servicios de transferencia de datos, *Web Broadcast* y *SlideShow*, que no contemplan la transmisión de contenido audiovisual sobre MOT.

Por otra parte, la norma DVB-H está basada en el estándar DVB-T, introduciendo algunas modificaciones para mejorar el rendimiento en terminales móviles y adoptando protocolos IP multicast para el transporte. La norma permite integrar servicios de DVB-H y DVB-T en una misma frecuencia, aprovechando así la infraestructura de los sistemas terrestres de difusión de televisión digital. Los estudios demuestran que DVB-H requiere, además, de infraestructura adicional para ofrecer cobertura suficiente a terminales móviles para los servicios de *streaming*. En Europa, los países que han completado el apagón analógico han dedicado parte de los recursos espectrales liberados (lo que se conoce como dividendo digital) a la implantación de servicios de DVB-H. Así, esta tecnología ya es comercial en países como Italia, Finlandia, Holanda o Suiza. Para los protocolos de transporte, es posible utilizar dos estándares bastante similares: DVB *IP DataCast* (IPDC) y OMA BCAST. En la actualidad se trabaja en la interoperabilidad de las dos normas. Ambas contemplan el uso de FLUTE para los servicios de transferencia de ficheros. No existe ninguna interfaz de programación estandarizada para DVB-H.

Las administraciones europeas, celosas de una posible fragmentación que pudiera debilitar el mercado interior de la televisión móvil, favorecieron la adopción de DVB-H a nivel europeo. Hoy en día, y dado que no se puede forzar a ningún país miembro a tomar una decisión de estas características, la Unión promueve la interoperabilidad entre los diferentes sistemas. Los principales inconvenientes de DVB-H frente a T-DMB son la disponibilidad de frecuencias (T-DMB utiliza menos ancho de banda y en una banda menos utilizada) y paradójicamente, su enorme capacidad (hasta 11Mbps, o más de 40 canales de televisión) que se considera excesiva en algunos mercados con poca demanda de difusión. Existe una variación de la norma, basada en el estándar DVB-S2 y denominada DVB-SH que utiliza transmisión por satélite y una red de transmisores terrestres para mejorar la recepción a móviles en interiores. Está previsto el lanzamiento de un satélite para ofrecer servicios de DVB-SH en Europa.



Tanto T-DMB como DVB-H/SH requieren de nuevas infraestructuras de comunicaciones para su implantación y es muy probable que su despliegue en los próximos años esté altamente condicionado por el desarrollo de la crisis financiera en la que nos encontramos. En la actualidad se está popularizando una nueva alternativa que no requiere de nueva infraestructura: la recepción de servicios de DVB-T en los móviles. Esta tercera alternativa es posible gracias a los avances que se han producido recientemente en los chipsets receptores de DVB-T de última generación, con menor consumo de energía y mayor sensibilidad, lo que posibilita su instalación en terminales móviles. Los problemas que presenta esta alternativa son claros: los formatos y los codecs (MPEG-2 vídeo y audio) que se utilizan en la televisión terrestre no son aptos para terminales móviles y los sistemas de TDT no se han planificado para su recepción en este tipo de terminales. De esta forma, no se puede garantizar una calidad de servicio aceptable en el área de cobertura. Una segunda generación de esta alternativa, en fase de desarrollo en la actualidad, utiliza codificación H.264 y FEC a nivel de transporte MPEG-TS para solventar estos problemas.

Por lo que respecta a España, el Ministerio de Tecnología Industria y Comercio lanzó en la primavera de 2007 una consulta sobre la Televisión Digital móvil. Como resultado de esta consulta, el Ministerio anunció en 2008 su voluntad de incluir en la planificación de los servicios de televisión digitales un múltiple para servicios de TDT a móviles sobre DVB-H. El despliegue de este múltiple se iniciaría una vez concluida la transición de los servicios de televisión a plataformas digitales, previsto para 2010.

## **II. MULTIPLEXACIÓN DE OBJETOS EN SISTEMAS DE DIFUSIÓN**

La optimización de la difusión de objetos en redes de comunicación es un problema ampliamente estudiado en la literatura. Esto es debido a que, tradicionalmente, el tráfico de las aplicaciones de transferencia de datos es intrínsecamente asimétrico: una gran cantidad de usuarios se conectan a un servidor de datos para descargar información. El volumen de los datos solicitado es sensiblemente mayor al necesario para señalar la solicitud. Dada esta naturaleza asimétrica del tráfico, la transmisión de documentos multimedia sobre redes de difusión se ha considerado como un mecanismo potencialmente eficiente en la transferencia de objetos a grandes audiencias. Otro estímulo para la investigación en este campo lo ha constituido la posibilidad de transmitir objetos multimedia sobre las redes digitales de difusión de contenidos.

La metodología más extendida en la bibliografía para analizar las prestaciones de la transmisión cíclica de objetos en redes de difusión fue planteada en [2], donde se formula que el tiempo de acceso a un objeto viene dado por:

| Símbolo   | Descripción   |
|-----------|---|
| $C$       | Colección de Objetos que se transmiten<br>( $C = \{O_1, O_2, \dots, O_M\}$ )  |
| $O_i$     | Objeto con índice $i$   |
| $T_{A_i}$ | Tiempo de Acceso, o tiempo transcurrido desde que se selecciona el objeto $i$ hasta que está disponible a nivel de aplicación (seg) |
| $T_{D_i}$ | Tiempo de Descarga (seg)  |
| $T_W$     | Tiempo de Espera, transcurrido desde que se selecciona el objeto hasta que está disponible para la descarga (seg)                   |
| $S_i$     | Tamaño del objeto (bits)  |
| $b_i$     | Velocidad de transmisión del objeto $i$   |
| $p_i$     | Probabilidad de acceso al objeto $i$  |

Tabla 1. Parámetros que definen el tiempo de acceso a los objetos

$$T_{A_i} = T_{D_i} + T_W = S_i / b_i + T_W \quad \forall O_i \in C \quad (1)$$

En [2] se estudia qué proporción del ancho de banda total  $b$  debe ser asignado a la transferencia de cada objeto,  $b_i$ , para minimizar el Tiempo de Acceso medio en el área de servicio. En este estudio se demuestra que si se conoce con qué probabilidad se va a acceder a los objetos, entonces se puede obtener una serie de valores de  $b_i$  que minimizan (1), como veremos a continuación, aunque basándonos en otro estudio. En [2], se divide el ancho de banda en  $N$  sub-tramas de difusión de capacidad  $b/N$  y se busca un mínimo en el tiempo de espera. El problema de este análisis es que, al dividir el ancho de banda total en sub-tramas, el tiempo de descarga de los objetos es directamente proporcional a  $N$ , por lo que no se garantiza que el mínimo del tiempo de espera se corresponda con el mínimo en el tiempo de acceso. Por eso, en este trabajo se elabora el método propuesto en [1], que utiliza la metodología planteada en [2], pero propone una distinción en el ancho de banda asignado a cada fichero en los dos términos de (1), ya que el ancho de banda utilizado en el tiempo de descarga se corresponde con el ancho de banda instantáneo, mientras que el ancho de banda utilizado por cada objeto en el tiempo de espera se corresponde con un promedio a largo plazo. [1] deduce que el mínimo del tiempo de descarga se obtiene cuando se asigna toda la capacidad disponible  $b$  para la descarga:

$$E[T_{A_i}] = T_{D_i} + E[T_W] = S_i / b + \frac{1}{2} \sum_{j=1}^M S_j / b_j \quad (2)$$

Donde se asume que el tiempo en que se solicita el objeto es una variable aleatoria uniformemente distribuida y que no hay interrupciones en la transmisión de los objetos debidas a errores en la transmisión o a la transmisión de fragmentos de otros objetos.

Para obtener el tiempo medio de acceso a todos los objetos en el área de servicio, se realiza un promedio en el tiempo de acceso de cada uno de los objetos en  $C$ . En el cálculo, se asume que los objetos no tienen la misma probabilidad de acceso en el área de servicio. Es decir, se asume que habrá objetos que se acceden más veces que otros. Este promedio se obtiene al descondicionar (2) con  $i$ :

$$E[T_A] = \sum_{i=1}^M S_i \cdot p_i / b + \frac{1}{2} \sum_{j=1}^M S_j \cdot p_j / b_j \quad (3)$$

El mínimo del tiempo de espera en (3) viene dado por  $B = \{b_1, b_2, \dots, b_M\} = \{\sqrt{S_1 p_1}, \sqrt{S_2 p_2}, \dots, \sqrt{S_M p_M}\}$ . Normalizando y adaptando el resultado a la capacidad disponible, se obtiene que la tasa a largo plazo que se debe asignar a cada objeto viene dada por:

$$b_i = \frac{\sqrt{S_i p_i} \cdot b}{\sum_{j=i}^M \sqrt{S_j p_j}} \quad (4)$$

Como se puede observar en (4), los objetos más populares tendrán una tasa a largo plazo mayor y por tanto, un ciclo de transmisión más corto. Lo mismo ocurre con los ficheros de mayor tamaño. Una vez que se conoce la tasa asignada a cada objeto en función de sus características, el problema de garantizar dicha tasa al multiplexar la transmisión de los distintos objetos en el tiempo es análogo al de control de tasas de las políticas de *fair-queuing* en las redes de conmutación de paquetes. Se pueden establecer analogías entre paquetes y objetos, y entre flujos y la transmisión cíclica de un objeto.

En general, al comparar los algoritmos de control de tasa en colas de paquetes, se utilizan dos parámetros: la precisión con la que el algoritmo ofrece la tasa requerida por cada flujo y el *jitter* o la variación en el retardo de los paquetes. En la transmisión cíclica de objetos, la precisión determinará el grado con el que se garantiza la media estadística a largo plazo de la tasa con la que se transmite un objeto. Si el algoritmo utilizado para determinar la secuencia de transmisión de los

objetos no consigue ajustar con precisión los valores obtenidos en (4), el tiempo medio de acceso en el área de servicio se desviará del mínimo, siempre que se cumpla que la probabilidad de acceso a cada objeto en el área de servicio es la misma que la utilizada para obtener las tasas. Por otra parte, el *jitter* estará ligado a la periodicidad del ciclo de transmisión de un objeto y, de manera indirecta, al consumo de batería en el entorno móvil: si ocurren errores en la transmisión, el móvil tendrá que esperar al siguiente ciclo para completar la descarga. Si la transmisión es aproximadamente periódica, el terminal podrá predecir el comienzo del siguiente ciclo y desconectar el receptor hasta entonces sin necesidad de una indicación explícita por parte del servidor acerca del próximo ciclo de transmisión.

### III. GESTIÓN DE CACHÉ Y PERSONALIZACIÓN

En el capítulo anterior se ha demostrado que el problema de ajustar la tasa de un objeto de acuerdo con su probabilidad de acceso en el canal de difusión se puede resolver analíticamente. Optimizar el tiempo de acceso a los objetos se reduce al problema de garantizar que la tasa obtenida para cada uno de los flujos se corresponda con la tasa obtenida a través del estudio analítico. Este problema también forma parte del estudio de las políticas de *fair queuing* en las redes de conmutación de paquetes. La transmisión de los objetos a estas tasas permite optimizar el uso de la capacidad en media, en base a un promedio de la popularidad de los objetos, pero no para un usuario en particular: las preferencias de un usuario, la popularidad relativa de los contenidos percibida por un cliente, no tienen por qué ajustarse a la estimación o la medida empírica de la probabilidad de acceso que establece el servidor. Si las preferencias específicas de un usuario difieren significativamente de las que estima el servidor, es muy probable que la experiencia percibida por ese usuario sea penalizada porque las tasas de los objetos a los que accede tienen una tasa menor en la transmisión.

Este efecto negativo ha sido ampliamente estudiado en [3], pero dado que es un estudio anterior a [1], se emplea un método diferente para ajustar la tasa de los objetos en base a su probabilidad de acceso, denominado *Broadcast Disks*. De los resultados de este estudio se desprende que, aunque en general, el método de ajustar la tasa de los objetos en base a su popularidad ofrece mejores resultados que el *flat broadcasting* (transmitir todos los objetos con la misma tasa), algunos usuarios se verán penalizados por la predicción de la popularidad que establece el servidor. En este estudio se demuestra que la técnica de *caching* (el almacenamiento en memoria de objetos en el carrusel antes de que el usuario solicite una descarga) mejora considerablemente el tiempo medio de acceso de los usuarios que son penalizados y, consecuentemente, el tiempo medio de acceso global.

Otra de las contribuciones de [3] es que presenta técnicas de *caching* adaptadas a entornos de difusión. Tradicionalmente, estas técnicas han sido establecidas y estudiadas para mejorar las

prestaciones del acceso a páginas *web*. Los algoritmos de gestión de caché en este entorno tienen como objetivo almacenar en caché las páginas que tienen una mayor probabilidad de acceso en el futuro más inmediato.

En [3] se demuestra que, en un sistema de difusión, el algoritmo de gestión de caché debe almacenar los objetos para los que la probabilidad de acceso local es significativamente mayor que la tasa de transmisión del objeto, ya que el valor de almacenar un objeto en caché es inversamente proporcional a su tasa de transmisión: si un objeto se transmite a una tasa mayor, la diferencia entre acceder al objeto desde la transmisión o desde memoria se reduce. Para evaluar la mejora introducida por las técnicas de *caching*, el estudio plantea un algoritmo denominado *Probability Inverse Frequency of broadcast* (PIX) que evalúa el ratio entre la probabilidad de acceso local y la frecuencia con la que un objeto se transmite para determinar el valor de cada objeto en memoria en las políticas de reemplazamiento en caché, de acuerdo con las consideraciones anteriores.

#### IV. ESTUDIO DE LA ARQUITECTURA DEL SERVICIO

*Digital Video Broadcasting – Handheld* (DVB-H) [4] es un sistema para el transporte eficiente de servicios multimedia a terminales móviles, a través de redes de difusión de terrestres. A nivel físico, el enlace está basado en la norma DVB-T (*Digital Video Broadcasting – Terrestrial*), en la que introduce ciertas modificaciones a nivel físico y a nivel de enlace de datos para mejorar las prestaciones del sistema en movilidad. Los parámetros de configuración de estos niveles se determinan en la fase de planificación de frecuencias de red, para garantizar la cobertura en la zona de servicio y establecen la capacidad total del sistema:

| Modulation | CR  | Intervalo de Guarda |        |        |        |
|------------|-----|---------------------|--------|--------|--------|
|            |     | 1/4                 | 1/8    | 1/16   | 1/32   |
| QPSK       | 1/2 | 4.976               | 5.529  | 5.855  | 6.032  |
|            | 2/3 | 6.635               | 7.373  | 7.806  | 8.043  |
|            | 3/4 | 7.465               | 8.294  | 8.782  | 9.048  |
|            | 5/6 | 8.294               | 9.216  | 9.758  | 10.053 |
|            | 7/8 | 8.709               | 9.676  | 10.246 | 10.556 |
| 16-QAM     | 1/2 | 9.953               | 11.059 | 11.709 | 12.064 |
|            | 2/3 | 13.271              | 14.745 | 15.612 | 16.086 |
|            | 3/4 | 14.929              | 16.588 | 17.564 | 18.096 |
|            | 5/6 | 16.588              | 18.431 | 19.516 | 20.107 |
| 64-QAM     | 7/8 | 17.418              | 19.353 | 20.491 | 21.112 |
|            | 1/2 | 14.929              | 16.588 | 17.564 | 18.096 |

|     |        |        |        |        |
|-----|--------|--------|--------|--------|
| 2/3 | 19.906 | 22.118 | 23.419 | 24.128 |
| 3/4 | 22.394 | 24.882 | 26.346 | 27.144 |
| 5/6 | 24.882 | 27.647 | 29.273 | 30.160 |
| 7/8 | 26.126 | 29.029 | 30.737 | 31.668 |

Tabla 2. Tasa de transmisión a nivel de enlace (Mbps) para un canal de 8MHz

A nivel de Enlace de Datos, DVB-H utiliza *Moving Picture Expert Group (MPEG) Transport Stream (TS)* [5]. Por tanto, toda la capacidad ofrecida por la capa física se utiliza para difundir una trama de MPEG-TS de tasa constante. Las especificaciones para la Información de Servicio [6] definen la semántica de los metadatos necesarios para incluir múltiples servicios en una trama de MPEG-TS. Para encapsular los datagramas IP en la trama de MPEG-TS, DVB-H utiliza una modificación del protocolo *Multi Protocol Encapsulation (MPE)* definido en [7]. MPE define un formato de paquete intermedio, denominado sección MPE, para transportar datagramas IP sobre MPEG-TS. Las modificaciones introducidas son compatibles hacia atrás con la especificación MPE original e introducen dos funcionalidades, que son opcionales: *Time Slicing*, para mejorar el consumo de batería y *MPE-FEC (Forward Error Correction)* para mejorar la recepción en terminales móviles.

En cuanto a los protocolos de transporte de contenidos, existen dos normas compatibles con el enlace DVB-H: *DVB-H Content Delivery Protocols (CDP)* [8] y *OMA BCAST Transport Protocols* [9]. Ambos estándares adoptan el uso de protocolos IP multicast para el transporte de dos tipos de servicio: servicios en tiempo real (*streaming*) y servicios de descarga de ficheros. Los servicios de *streaming* utilizan el protocolo RTP y los servicios de descarga de ficheros utilizan FLUTE [10]. Además, las normas definen un servicio de descarga de ficheros, denominado *Electronic Service Guide (ESG)*, *OMA ESG* [11] y *DVB ESG* [12], y que sirve para proporcionar a los terminales información de acceso a los servicios y a los usuarios información sobre los servicios y sobre los contenidos. En las siguientes secciones se explica el funcionamiento de FLUTE y la estructura de la ESG para los servicios de descarga de ficheros.

#### IV.1.EL PROTOCOLO FLUTE

El *Internet Engineering Task Force (IETF)* define FLUTE en la RFC 3926 [10] como un protocolo unidireccional fiable y escalable para la transmisión masiva de ficheros en *Internet* sobre IP multicast. Para desempeñar esta función, tal y como se indica en [13], un protocolo debe soportar la segmentación del paquete a nivel de transporte, el envío paralelo de múltiples ficheros, la integración de mecanismos de corrección de errores (FEC), corrección repetitiva, control de la congestión y señalización de las propiedades de los ficheros. Todas estas funciones excepto la última las proporciona el protocolo sobre el que se construye FLUTE, *Application Layer Control*

(ALC) [14], que es un protocolo para la transmisión masiva y escalable de objetos a través de *Internet*.

FLUTE utiliza el concepto de Bloques Constructivos para el transporte escalable de transferencias masivas de datos sobre redes multicast, descrito en [15]. ALC se define como un protocolo para la entrega escalable y fiable de contenido sobre redes multicast y consiste en la integración de tres Bloques Constructivos, *Layer Control Transport* (LCT) [16] para el transporte, un Bloque de control de la congestión y un bloque de corrección de errores FEC [17].

El bloque LCT proporciona soporte a nivel de transporte a protocolos de *streaming* o de transferencia de objetos diseñados en base al concepto de Bloques Constructivos. Ofrece una cabecera flexible y extensible, para adaptarse a diferentes casos de uso. La cabecera incluye un identificador de sesión, *Transport Session Indicator*, (TSI) y un identificador de objeto (TOI), único para cada objeto que se transmite en una sesión.

El bloque de Control de la Congestión añade la información de control necesaria para implementar un algoritmo de control de la congestión sobre un flujo unidireccional en *Internet*. Como el uso de un algoritmo de estas características no tiene sentido alguno en una red de difusión inalámbrica como DVB-H, este bloque no tiene ninguna relevancia en el contexto de este trabajo.

Por último, el bloque FEC [17] define los parámetros de señalización necesarios para aplicar FEC a nivel de aplicación (AL-FEC), con el objetivo de aumentar la fiabilidad en el transporte de datos. Estos parámetros son el Identificador de Codificación FEC, para reconocer el código que se aplica, la Información de Transmisión de Objeto FEC, que proporciona los parámetros de configuración del código FEC utilizado y un Identificador de Contenido FEC, necesario para identificar los símbolos codificados dentro de un paquete.

FLUTE utiliza ALC a nivel de transporte y define dos tipos de objetos, los ficheros propiamente dichos y la *File Delivery Table* (FDT), que proporciona una descripción de las propiedades de los ficheros. La FDT proporciona una relación entre las propiedades del fichero y el TOI que se utiliza para su transporte. La FDT se envía a través de Instancias FDT, escritas en lenguaje XML, con TOI igual a 0.

Para unirse a una sesión de FLUTE, un cliente debe conocer los parámetros de transporte de la sesión, que son la dirección IP multicast de destino, el puerto UDP utilizado y el indicador de sesión TSI de los paquetes ALC. En DVB-H esta información se proporciona a través de la ESG en formato *Session Description Protocol* (SDP) [18] definido en [10] para este propósito.

El Apéndice A contiene una descripción del formato de la cabecera de ALC. El proceso para transmitir un fichero sobre DVB-H basado en el protocolo FLUTE se divide en los siguientes pasos:

1. El fichero puede, opcionalmente, comprimirse con GZIP para reducir el tamaño del mismo antes de iniciar la sesión de transferencia. Si se utiliza compresión para la transmisión de un fichero, se debe señalar en la entrada correspondiente de la FDT.
2. El fichero comprimido se divide en  $Z$  bloques, denominados bloques fuente. Los bloques fuente se dividen a su vez en  $N$  sub-bloques iguales, de tamaño configurable. Los sub-bloques son la unidad de decodificación y de su tamaño depende la memoria requerida por las tareas de decodificación en el receptor.
3. Los sub-bloques se dividen a su vez en  $K$  sub-símbolos de tamaño  $T''$ . El símbolo de codificación  $i$  es la concatenación del sub-símbolo con índice  $i$  de cada uno de los sub-bloques. Los símbolos de codificación tienen un tamaño  $T$ . Los símbolos de codificación que se corresponden con el fichero original se denominan símbolos fuente. Señalar que el valor de  $K$  no es necesariamente el mismo para todos los sub-bloques que constituyen un fichero y que el valor de  $T''$  no tiene por qué ser necesariamente el mismo en todos los sub-símbolos que dan lugar a un símbolo de codificación.
4. Sobre cada bloque, se aplica el algoritmo de codificación AL-FEC, que debe ser sistemático para garantizar la recepción por parte de terminales que no implementen el decodificador FEC. Así, todos los receptores deben ser capaces de reconstruir un fichero a partir de los símbolos. Es decir, aunque no se aplique AL-FEC, se debe los ficheros se dividen en bloques y símbolos para el transporte.
5. Cada paquete transporta o bien únicamente símbolos fuente o símbolos de reparación. El *payload* de un paquete de FLUTE contiene  $G$  símbolos de codificación pertenecientes al mismo bloque.
6. El uso de AL-FEC y MPE-FEC es opcional y mutuamente excluyente. Si no se utiliza AL-FEC, el fichero se sigue reconstruyendo a partir de símbolos de codificación, pero en este caso, todos los símbolos de codificación se corresponden con símbolos fuente. El único algoritmo de FEC adoptado por DVB-H son los códigos Raptors [Raptors]. Los códigos Raptors pertenecen a la categoría de *Fountain Codecs* y tienen características similares a la codificación ideal [17].

El tamaño de un objeto FLUTE originado a partir de un programa dependerá de los parámetros utilizados para dividir el objeto en bloques, los bloques en símbolos, para encapsular los símbolos en paquetes ALC y de la configuración de la cabeceras de ALC, UDP y IP multicast. Este proceso viene dado por las siguientes expresiones:



| Símbolo            | Descripción   |
|--------------------|---|
| $D$                | Duración del programa (segundos)  |
| $C_{cod}$          | Coefficiente obtenido al generar un fichero MP4 a partir del programa (bps)   |
| $F$                | Tamaño del fichero a transmitir (bytes)   |
| $T$                | Tamaño aproximado de los símbolos de codificación (bytes)   |
| $Z$                | Número de bloques en que se divide el fichero   |
| $G_{MAX}$          | Máximo número de símbolos por paquete IP  |
| $K_T$              | Número de símbolos de codificación en un fichero  |
| $K_{MIN}, K_{MAX}$ | Número mínimo y máximo de símbolos por bloque   |
| $A$                | Parámetro de alineamiento de símbolo (bytes)  |
| $P$                | Payload máximo en un paquete (bytes)  |
| $Partition(I, J)$  | Función que toma dos valores enteros I y J como entrada y devuelve $I_L = \text{ceil}(I/J)$ , $I_S = \text{floor}(I/J)$ , $J_L = I - I_S$ y $J_S = J - J_L$ |
| $S_{HEADER}$       | Tamaño de las cabeceras IP, UDP, y ALC  |

Tabla 3. Parámetros de la encapsulación y la paquetización de objetos FLUTE

$$F = D \cdot C_{cod} / 8 \quad (5)$$

$$G = \min\{\text{ceil}(P \cdot K_{MIN} / F), P / A, G_{MAX}\} \quad (6)$$

$$T = \text{floor}(P / (A \cdot G)) \cdot A \quad (7)$$

$$K_T = \text{ceil}(F, T) \quad (8)$$

$$(K_L, K_S, Z_L, Z_S) = \text{partition}[K_T, Z] \quad (9)$$

$$S = Z_L \cdot (\text{floor}(K_L / G) \cdot (G \cdot T + S_{HEADER}) + ((\text{rem}(K_L, G) \cdot G) \cdot T + S_{HEADER})) + Z_S \cdot (\text{floor}(K_S / G) \cdot (G \cdot T + S_{HEADER}) + ((\text{rem}(K_S, G) \cdot G) \cdot T + S_{HEADER})) \quad (10)$$

La función  $Partition(I, J)$  divide un objeto de tamaño  $I$  en  $J$  bloques de tamaño aproximadamente iguales, concretamente  $I_L$  bloques de tamaño  $J_L$  y  $I_S$  bloques de tamaño  $J_S$ .

Así, el fichero se dividirá en  $Z = Z_L + Z_S$  bloques, de tamaño  $K_L \cdot T$  para los  $Z_L$  primeros y  $K_S \cdot T$  para los  $Z_S$  restantes. Los bloques se transmitirán en paquetes con  $G$  símbolos por paquete, a excepción del último paquete de cada bloque, que transportará un número menor de símbolos (exactamente  $\text{rem}(K_L, G) \cdot G$  para los  $Z_L$  primeros bloques y  $\text{rem}(K_S, G) \cdot G$  para los últimos). Este proceso determinará el tamaño de cada objeto FLUTE que se debe utilizar en (3) para determinar los tiempos medios de acceso a los ficheros.

## IV.2.LA GUÍA DE SERVICIO OMA BCAST SERVICE GUIDE

La Guía Electrónica de Servicios, *Electronic Service Guide* (ESG), es un servicio de transferencia de ficheros utilizado para proporcionar a terminales y usuarios una descripción de los servicios disponibles sobre una plataforma IP. Se puede considerar que la ESG es el equivalente a las Guías Electrónicas de Programación (*Electronic Program Guides*, EPGs) que se ofrecen en los servicios de Televisión Digital basados en las normas DVB, como la TDT.

Como se ha comentado con anterioridad, existen dos estándares que definen servicios de ESG compatibles con la norma DVB-H. Estos son, por un lado OMA BCAST *Service Guide* [11] y DVB IPDC *Service Guide* [12]. Ambos estándares comparten más similitudes que diferencias. Por un lado, ambos consisten en servicios de transferencias de ficheros sobre FLUTE, a diferencia de las EPGs de DVB, constituidas por secciones de MPEG-TS. Las dos normas aplican modelos de datos basados en el estándar XML y gestionan la adquisición y la actualización de los datos por métodos casi idénticos. La principal diferencia entre las dos normas está relacionada con la funcionalidad: la norma OMA BCAST está mucho más orientada a su uso sobre plataformas híbridas, que integran un canal de interacción, y también introduce funciones de interactividad que no están presentes en la norma de DVB. La norma OMA también es más flexible en cuanto a los tipos de servicios ofrecidos.

La norma OMA contempla el tipo de servicio *Cachecasting*. El modelo de datos especifica que la descripción XML de una instancia de la ESG se divida en fragmentos que se transmiten independientemente. Cada tipo de fragmento contiene información sobre diferentes aspectos del servicio. Los tipos de fragmentos más importantes en el contexto de este trabajo son:

- *Service*: los fragmentos del tipo *Service* contienen información general sobre el servicio, como el nombre y el tipo de servicio. Es el fragmento alrededor del cual se organizan todos los fragmentos que contienen información sobre un mismo servicio. Los demás fragmentos que componen la guía de este servicio incluirán una referencia a este fragmento.

- *Schedule*: indican los intervalos de tiempo en que los programas o contenidos están disponibles en el canal de difusión. El objeto de los fragmentos *Schedule* es informar a los terminales acerca de la disponibilidad de los contenidos.
- *Content*: los fragmentos de tipo Content proporcionan información descriptiva acerca de los contenidos, como información de género, de control parental, del horario de programación (para mostrar a los usuarios) o de idioma.
- *Access*: proporciona a los terminales la información necesaria para la adquisición de los contenidos, generalmente referenciando a un fragmento de una SDP. La SDP puede indicar que los contenidos están disponibles en difusión (FLUTE o RTP) o bajo demanda (HTTP RTSP).

En una Guía de Servicios para un servicio de Cachecasting, el atributo *type* de los fragmentos de *Service* debe ser igual a 4, para indicar que es un servicio de transferencia de ficheros. Para cada uno de los ficheros que se ofrecen a los usuarios, la ESG debe incluir un fragmento de tipo *Content*. El atributo *weight* indica a los terminales el orden de presentación de los ficheros. La información de género, idioma y control parental en este fragmento se utiliza para implementar las funciones de personalización: los terminales asignarán prioridad a la descarga de ficheros cuyas descripciones se correspondan con las preferencias de los usuarios. Los fragmentos de *Schedule* que hagan referencia al servicio y al fragmento *Content* de un fichero contendrán el atributo *ContentLocation*, que se corresponde con el elemento Content-Location del fragmento de la FDT que hace referencia al mismo fichero en la sesión de FLUTE. El fragmento de *Schedule* contendrá una o varias instancias del elemento *DistributionWindow* que proporciona información a los terminales sobre los intervalos de tiempo en que el fichero está presente en el carrusel. Esta información puede ser utilizada por los terminales para apagar el receptor de DVB-H en los intervalos de tiempo en que la sesión de FLUTE se utilice para transmitir otros ficheros. Los fragmentos de *Access* vinculados a la descarga de este fichero contendrán una referencia al servicio, otra al fragmento de *Schedule* correspondiente y alternativamente, otra referencia al elemento *DistributionWindow*. El fragmento referenciará una SDP con la información sobre la sesión de FLUTE (IP destino, puerto, TSI, etc.).

El siguiente esquema muestra la estructura básica de una Guía de Servicios para un servicio de transferencia de ficheros:

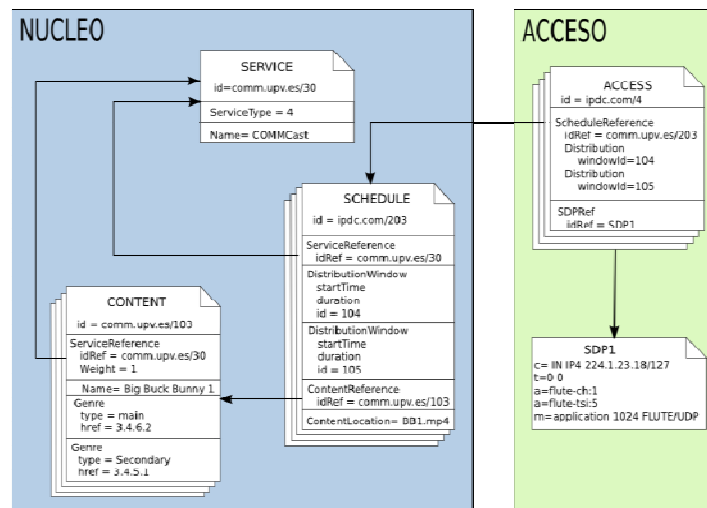


fig. 1. Estructura de la Guía Electrónica de Servicios para un Servicio de CacheCasting

#### IV.1. IMPLEMENTACIÓN DE ALGORITMOS DE REEMPLAZAMIENTO EN CACHÉ BASADOS EN PIX

La implementación de algoritmos de reemplazamiento en caché basados en *PIX* no es inmediata, debido a que no es posible predecir con exactitud la probabilidad de acceso a los contenidos por parte del usuario, pero es posible implementar algoritmos sub-óptimos, en la medida en que se pueda obtener una estimación de la misma. En el caso bajo estudio, la probabilidad de acceso está ligada a las preferencias del usuario sobre los contenidos. Las herramientas software destinadas a proporcionar una estimación de las preferencias de los usuarios se denominan herramientas de recomendación [20]. Estas herramientas se agrupan en dos categorías, dependiendo del tipo de información que utilicen para predecir las preferencias de los usuarios. Hay recomendadores basados en descripciones de los contenidos (*content-based recommenders*) y el segundo tipo utiliza valoraciones realizadas por otros usuarios (*collaborative filtering*). Dado que la ESG proporciona descripciones de los contenidos y que DVB-H es un sistema unidireccional, la primera categoría es a priori la más apropiada en este estudio. A continuación se presentan algunas características generales de los recomendadores basados en contenido y guías de implementación para la utilización de estas herramientas en algoritmos de gestión de caché *PIX* en aplicaciones de *CacheCasting* basadas en la norma OMA BCAST.

Los recomendadores basados en contenidos necesitan conocer las preferencias del usuario, bien en base a una descripción explícita proporcionada por los usuarios, o bien en base a un análisis del historial de los accesos a contenidos del cliente. Una vez que se dispone de esta estimación, las herramientas deben implementar un método heurístico para comparar los contenidos y las preferencias del usuario y estimar la utilidad de cada uno de los contenidos para el usuario. Así, un recomendador debe integrar una función  $u$  que mida la utilidad de cada del catálogo  $C$  y obtenga una clasificación,  $R$ , de los contenidos en base a su utilidad ( $u : C \times R$ ). La clasificación  $R$

(de *ratings*) es una serie ordenada de valores enteros o reales positivos en un intervalo acotado.

**Para utilizar un recomendador en la gestión de caché, asumimos que la probabilidad de acceso futuro de un contenido es directamente proporcional a su utilidad:  $R \propto P$ .**

El método más inmediato para implementar la función  $u$  consiste en la comparación de una descripción explícita del perfil del usuario  $c$  y de lo cada programa  $s$  en un espacio semántico en el que cada dimensión representa el peso de una palabra clave. En base a este método, una descripción de un contenido  $s$  se puede representar como un vector de pesos  $\vec{w}_s = (w_{s1}, w_{s2}, \dots, w_{sk})$ , donde  $w_{si}$  denota la importancia de la palabra clave  $k_i$  en el contenido. El perfil del usuario vendrá dado por el vector  $\vec{w}_c = (w_{c1}, w_{c2}, \dots, w_{cK})$  y la utilidad del contenido es directamente proporcional a la distancia entre los dos vectores en el espacio semántico:

$$u(c,s) = \cos(\vec{w}_s, \vec{w}_c) = \frac{\sum_{i=1}^K w_{i,c} w_{i,s}}{\sqrt{\sum_{i=1}^K w_{i,c}^2} \cdot \sqrt{\sum_{i=1}^K w_{i,s}^2}} \quad (11)$$

Esta subcategoría de métodos se denominan recomendadores heurísticos, dado que utilizan una fórmula, como la del coseno, para determinar la similitud entre el perfil del usuario y la descripción de un contenido. Algunas mejoras que se pueden introducir para mejorar las prestaciones de este tipo de recomendadores pueden ser, por un lado, computar la distancia entre el perfil y el usuario en varios espacios semánticos diferentes (ejemplo, un espacio pueden ser descripciones de género, otro espacio puede constituirse por la sinopsis, otro por los directores, etc.) y luego realizar un promedio. Así se puede aprovechar más información para realizar la comparación. Otra opción es utilizar diccionarios jerárquicos y representar las palabras claves en una estructura de árbol. La distancia entre dos términos se computa como la distancia entre dos nodos en el árbol, en lugar de distancias euclídeas en un espacio multidimensional. Así se tiene en cuenta la similitud semántica entre las palabras clave utilizadas. Por último, para prescindir de una descripción explícita del perfil del usuario,  $\vec{w}_c$ , se puede estimar la utilidad  $u(c,s)$  a partir de las utilidades  $u(c,s_j)$  asignadas por un usuario  $c$  a contenidos  $s_j \in S$  consumidos con anterioridad y que son similares a  $s$ .

Otra subcategoría entre los recomendadores son los basados en modelo, que utilizan modelos estadísticos de la información disponible para clasificar los contenidos. Los clasificadores Bayesianos (*Bayesian classifiers*) se basan en este principio para clasificar la utilidad de los contenidos en una serie de categorías (Por ejemplo,  $C_i \in [C_1, C_2, C_3, C_4, C_5] = [\star, \star\star, \star\star\star, \star\star\star\star, \star\star\star\star\star]$ ), que se utiliza en la mayoría de portales de *streaming* de vídeo en

*Internet*). Los clasificadores Bayesianos realizan la categorización en base a la probabilidad de que un contenido  $\dot{w}_s = (w_{s1}, w_{s2}, \dots, w_{sk})$  pertenezca a una categoría  $C_i$  asumiendo que:

$$P(C_i | \dot{w}_s) = P(C_i | k_{1,s} \& k_{2,s} \& \dots \& k_{N,s}) \quad (12)$$

donde  $k_{1,s}, k_{2,s}, \dots, k_{N,s}$  es el conjunto de palabras clave asignado al contenido  $s$ . Por el teorema de Bayes, la probabilidad anterior es directamente proporcional a:

$$P(C_i | \dot{w}_s) \propto P(C_i) \prod_x P(k_{x,s} | C_i) \quad (13)$$

Las dos probabilidades en el segundo término de (13) se obtienen de un análisis del historial del servicio:  $P(C_i)$  es la probabilidad de que un contenido  $s_j \in S$  pertenezca a la categoría  $C_i$ .  $P(k_{x,s} | C_i)$  es la probabilidad de que la palabra clave  $k_{x,s}$  ocurra en los contenidos dentro de la categoría  $C_i$ .

Para los dos métodos presentados, la gestión de caché debe , i) después de procesar la ESG de cada contenido en el carrusel, ii) obtener una descripción de los contenidos y iii) comparar dicha descripción con un modelo de las preferencias del usuario para iv) derivar una estimación de la probabilidad de acceso para cada *ítem* que sea utilizada en el algoritmo de gestión de caché basado en *PIX*.

Como se ha explicado con anterioridad, la Guía de un Servicio de *Cachecasting* contiene dos tipos de descripciones para cada programa en el carrusel de datos, una descripción proporcionada en un lenguaje controlado y una sinopsis o descripción textual en un lenguaje libre (normalmente dirigida al usuario final). Para la descripción en lenguaje controlado, la guía de OMA BCAST permite incluir descripciones de género (en formato *TV-Anytime* [21] o en formato *Moving Image Genre-Form*). Además de la información de género, la ESG admite descripciones sobre el público objetivo al que va destinado el programa, de nuevo en formato *TV-Anytime* o en un lenguaje definido por el operador. En este punto, asumimos que se utilizan las descripciones de género en formato *TV-Anytime*, el resto de consideraciones serían equivalentes si se utiliza información sobre el público objetivo, si se procesa la sinopsis para obtener palabras clave o si se realiza un promedio en los tres espacios semánticos. El resto de esta sección se dedica a proporcionar algunos detalles sobre la implementación de los algoritmos de recomendación en el contexto de la aplicación.

Al utilizar *TV-Anytime*, las palabras clave utilizadas para la comparación forman parte de un diccionario jerárquico acotado, por lo que no se requiere de ningún procesamiento de los contenidos en el cliente para obtener las palabras clave. En principio, el espacio semántico en el que realizar las

comparaciones de los contenidos y el perfil de usuario lo constituye la clasificación de contenidos, *ContentCS*, especificada en *TV-Anytime*, que es una clasificación jerárquica en múltiples niveles:

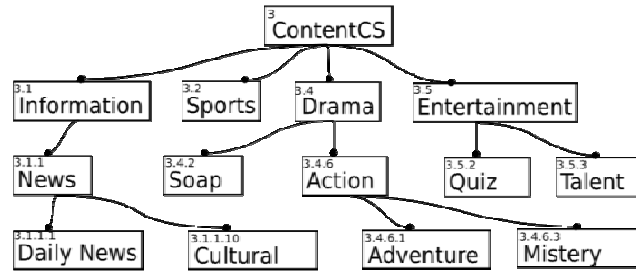


fig. 2. Estructura de los elementos *ContentCS* de *TV-Anytime*

Los niveles de jerarquía permiten que los terminales puedan determinar el género con mayor o menor precisión, dependiendo de sus capacidades. Por ejemplo, para una marca de género de valor 3.1.1.1 un terminal que soporta hasta el nivel 4 determinará que es un programa de noticias diario, mientras que un terminal que soporta el nivel 2, sólo será capaz de determinar que el programa es de carácter informativo. El estándar de OMA BCAST indica que los terminales pueden soportar hasta el nivel cuatro y se descarta el resto de niveles.

#### IV.1.1. IMPLEMENTACIÓN A PARTIR DE UN RECOMENDADOR HEURÍSTICO

Un recomendador heurístico debe, primero, obtener el peso de los términos presentes en la descripción de género de la ESG para el contenido  $s_j$ . Para computar el peso de cada una de las marcas de género, el método más básico se denomina TF-IDF (*Term Frequency – Inverse Document frequency*) [22]. Asumamos que la marca de género  $k_i$  aparece en el contenido  $s_j$  un número  $f_{i,j}$  de veces.  $M$  es el tamaño del catálogo y  $m_i$  es el número de veces que la marca  $k_i$  aparece en todo el catálogo de contenidos. El peso  $w_{ji}$  viene dado por:

$$w_{ji} = \frac{f_{i,j}}{\max_z f_{z,j}} \cdot \log \frac{M}{m_i} \quad (14)$$

Para calcular el peso de la marca de género  $k_i$  en el perfil de un usuario  $c$ , la aplicación puede solicitar al usuario una declaración implícita sobre sus preferencias hacia los contenidos que vengán marcados con  $k_i$ . Para reducir el número de opciones sobre las que el usuario tiene que manifestar su interés, la aplicación puede preguntar al usuario solamente acerca de sus preferencias sobre los géneros en el primer nivel de *ContentCS* (información, deportes, entretenimiento, etc.). Cada uno de los géneros descendientes a niveles inferiores heredaría el peso asignado por el usuario al género de primer nivel. Los pesos asignados a cada uno de las marcas por el usuario  $c$

determinarían el vector que representa el perfil de usuario  $\vec{w}_c = (w_{c1}, w_{c2}, \dots, w_{cK})$ . Una vez que se tienen los dos vectores, la utilidad del contenido se puede derivar por una función heurística como el coseno. El coseno proporciona un valor en el intervalo  $[0,1]$  por lo que se puede utilizar directamente como una estimación de la probabilidad de acceso futuro al contenido en un algoritmo de gestión de caché basado en *PIX*. Para actualizar el perfil del usuario en base al consumo, la aplicación puede cambiar el peso establecidos en base al historial de uso de los contenidos.

Con esta alternativa, el peso  $w_{ci}$  se modifica dependiendo del uso que haga un usuario de los contenidos que vienen descritos por la marca  $k_i$ . Por ejemplo, si un usuario visualiza un contenido con una marca  $k_i$ , entonces se incrementa  $w_{ci}$ . Si por el contrario, el usuario abandona el servicio mientras se está visualizando un contenido con la misma marca, entonces  $w_{ci}$  disminuye.

#### IV.1.2. IMPLEMENTACIÓN A PARTIR DE UN RECOMENDADOR BAYESIANO

En el contexto de la implementación de un algoritmo de gestión de caché basado en un clasificador Bayesiano, cada categoría  $C_i$  se debe corresponder a un rango de la probabilidad de acceso local (por ejemplo  $K = \{C_1, C_2, C_3, C_4\} = \{P \in [0, 0.25], P \in [0.25, 0.5], P \in [0.5, 0.75], P \in [0.75, 1]\}$ ). Para obtener la probabilidad de acceso local de un contenido presente en la ESG, la aplicación debe estimar la probabilidad de que dicho contenido pertenezca a cada una de las categorías  $P(C_i | \vec{w}_s) \forall C_i \in K$  de acuerdo con (13). Para estimar los dos términos de (13) se utiliza información del historial y se asume que  $P(C_i)$  es proporcional a la probabilidad de que un contenido del historial de contenidos consumidos ( $s \in H$ ) pertenezca a la categoría  $C_i$ . De igual manera, para estimar  $P(k_{x,s} | C_i)$ , la probabilidad de que se dé la marca de género  $k_x$  en la categoría  $C_i$ , se debe analizar el historial de marcas de género asignadas a contenidos que fueron clasificados bajo  $C_i$  por el usuario.

De nuevo, se necesita que el usuario introduzca información sobre sus preferencias en la aplicación para poder obtener un modelo de sus preferencias, pero en este caso, la información es a posteriori, es decir, una vez visualizado el contenido, el usuario debe clasificar el contenido en una de las categorías  $C_i$ . Los hábitos de consumo se pueden utilizar para sustituir la información proporcionada implícitamente por el usuario. En el ejemplo anterior, el sistema podría asignar  $C_4$  a los contenidos

### V. MÉTODOS Y RESULTADOS

#### V.1. EVALUACIÓN ANÁLITICA DE LA MULTIPLEXACIÓN DE OBJETOS

Los primeros resultados de este estudio buscan demostrar que es posible proporcionar niveles de Calidad de Experiencia aceptable bajo las condiciones impuestas por el sistema. Para ello, se han



realizado simulaciones para evaluar el compromiso entre el número de ficheros ofrecidos y el Tiempo de Acceso Medio observado en (3). Para establecer esta relación se debe modelar la popularidad de los objetos, el tamaño de los mismos y la tasa asignada al servicio.

Para modelar la popularidad de los contenidos se ha utilizado la función ZIPF [23], que viene dada por la siguiente expresión:

$$p_i = Cte \cdot \frac{1}{i^\alpha} \quad (15)$$

Es fácil observar en (4) que el valor de  $Cte$  no afecta al valor de las tasas a largo plazo para cada fichero, dado que al normalizar la tasa de transmisión para ajustar a la capacidad disponible, esta constante desaparece, por lo que el único parámetro que queda por definir en (15) es  $\alpha$ , que determina cómo de rápido disminuye la popularidad entre los diferentes programas. Sin embargo, el valor de  $Cte$  sí afecta al valor del tiempo de espera que se obtiene en (3). Para la evaluación de este límite, se ha utilizado la siguiente metodología: El parámetro  $\alpha$  se estima a partir de estudios anteriores a este trabajo. Para un valor de  $\alpha$ , se obtiene  $\bar{p}_i = \frac{1}{i^\alpha}$ . Para computar (3) se aplica  $p_i = \bar{p}_i / \sum_i \bar{p}_i$ . Así se consigue que la suma de las probabilidades de acceso sea igual a 1. Las diferentes pruebas que se han realizado en otros estudios para determinar  $\alpha$  han demostrado que se obtienen valores muy diferentes dependiendo del contexto. Por ejemplo en [24] se ha obtenido un valor de 0.5 para contenido audiovisual, inferior a los valores obtenidos para caracterizar el acceso a contenido web, que oscilan entre 0.6 y 0.8 [25]. En este estudio se han utilizado valores que oscilan en el intervalo  $\alpha \in [0.5, 0.8]$ .

Para modelar la longitud de los archivos, se ha optado por una aproximación similar. De las pruebas de usuario consultadas se desprende que la duración media de los programas de TV móvil debe rondar los dos minutos, por la duración media de las sesiones que se han medido. Sin embargo, no se han encontrado estadísticas que modelen la duración de los programas. En las simulaciones, se han utilizado tres modelos estadísticos diferentes. El primero, se han utilizado ficheros de tamaño constante, de duración media igual a dos minutos. En el segundo, la duración de los programas se ha modelado como una variable estadística uniformemente distribuida entre 30 y 300 segundos. En el tercero se ha utilizado una variable aleatoria de distribución exponencial de media igual a 120 segundos. Hay que tener en cuenta que la producción de contenidos a móviles es un mercado emergente y que existen pocos estudios estadísticos sobre la duración de los programas. Estos valores han sido establecidos de acuerdo con la información obtenida de las diferentes pruebas de usuario en torno a la TV móvil [26]. Para modelar la codificación se ha obtenido una aproximación pesimista, asumiendo que se realiza codificación a una tasa constante de aproximadamente 210kbps. Este valor se obtiene para generar una tasa a nivel de IP de valor

256kbps. Este valor se corresponde con la tasa definida para los servicios de *streaming* sobre redes 3G y también, como el ancho de banda típico de un canal de TV sobre DVB-H, por lo que nos permitirá comparar el servicio con servicios de *streaming*, sin cuantificar posibles ganancias de codificación.

Una vez que se conoce el tamaño de los ficheros, se aplica (10) para obtener el tamaño del objeto FLUTE que se debe transmitir. Los parámetros necesarios para modelar la paquetización de objetos FLUTE son iguales a los recomendados en el apéndice C del estándar [8]: ( $W = 256KB, A=4, G_{max}=10, K_{min}=1KB$  y  $K_{max}=2KB$ ). El tamaño máximo de paquete IP se ha fijado a 1544 bytes, también de acuerdo con el valor recomendado para el *Maximum Transport Unit* (MTU) en redes DVB-H. Estas consideraciones son extensibles al resto de simulaciones y medidas.

Por último, se van a obtener comparaciones entre tres algoritmos distintos para multiplexar los objetos y conformar un flujo de salida con los ficheros en una secuencia determinada. El primero, lo denominaremos *Fluid Fair Queuing* (FFQ). En las simulaciones para este método, computaremos (3) y obtendremos los tiempos medios de espera a través de este límite sin tener en cuenta ningún aspecto relacionado con la implementación. El segundo método, denominado *Round Robin* (RR), consistirá en la transmisión secuencial de cada uno de los programas sin que ninguno se repita en un ciclo (es decir, todos los archivos tienen la misma tasa a largo plazo). En la siguiente sección se utilizará otro método denominado *Virtual Clock* Modificado (MVC).

Para facilitar la comparación de los resultados, utilizaremos la siguiente nomenclatura que expresa el tiempo medio de acceso a los ficheros (OAT tal y como se muestra en (3)) para una configuración de los parámetros de la simulación:

$$OAT[A, P, S] = OAT[\{RR, FFQ, MVT\}, ZIPF(\alpha), \{constant, uniform, exponential\}] \quad (16)$$

Por ejemplo, el tiempo de acceso de una simulación con RR, una distribución de probabilidades ZIPF con  $\alpha = 0$  y duraciones de programas que siguen una distribución de probabilidad uniforme, se expresaría como  $OAT[RR, ZIPF(0), uniform]$ .

En la siguiente figura se muestra la evaluación del límite presentado en (3) obtenido a través de esta metodología, para diversas combinaciones de la popularidad de los contenidos y su duración:

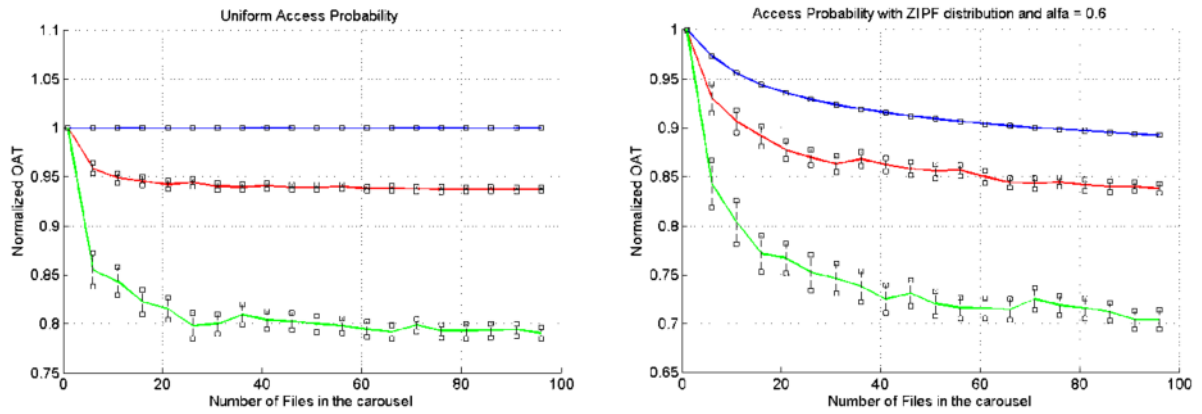


fig. 3. Tiempos Medios de Acceso para diferentes configuraciones de la multiplexación de objetos.

En el eje y se muestra el tiempo de acceso obtenido en (3) normalizado por el tiempo de acceso que se obtiene al transmitir los objetos con RR. Es decir, el eje y representa  $OAT[FFQ, ZIPF(\alpha), 'dist'] / OAT[RR, ZIPF(\alpha), 'dist']$  y se muestran los tiempos de accesos obtenidos para diferentes configuraciones de la transmisión con FFQ. En el eje de ordenadas se muestra el número de ficheros en el carrusel. La figura muestra los valores obtenidos para los tres modelos estadísticos del tamaño de los ficheros (azul = tamaño constante, rojo distribución uniforme y verde distribución exponencial). En la izquierda se muestran los resultados para una probabilidad de acceso uniforme ( $\alpha = 0$ ) y en la derecha para una distribución ZIPF con  $\alpha = 0.6$ . Los valores mostrados se corresponden con un nivel de confianza del 98%.

Para observar el efecto de la popularidad de los programas en el tiempo de acceso medio, en la siguiente gráfica se muestra el tiempo de acceso medio a los programas en minutos para diferentes valores del parámetro  $\alpha$  (azul 0.5, rojo 0.6, verde 0.7 y Amarillo 0.8) frente a la duración media de la oferta de contenidos en el carrusel, que se ha aproximado por el producto entre la duración media de los programas multiplicada por el número de ficheros en el carrusel. Los valores se muestran para ficheros de tamaño constante (línea continua) y con una distribución de tamaños exponencial (línea discontinua). De nuevo, se muestran valores con un margen de confianza del 98%

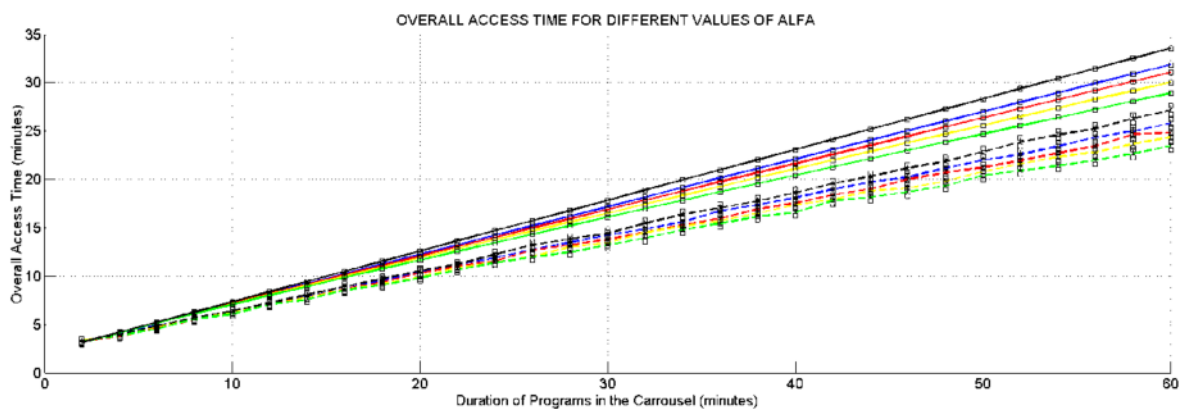


fig. 4. Tiempo medio de acceso para diferentes valores de  $\alpha$

## V.2.EVALUACIÓN DE LAS PRESTACIONES DEL ALGORITMO VIRTUAL CLOCK MODIFICADO

Para implementar la multiplexación de objetos se ha aplicado el algoritmo de *Virtual Clock* adaptado a entornos de difusión propuesto en [1] que se corresponde con el siguiente pseudocódigo:

```

Fase1:Inicialización
for i =1 to M
    object (i) .delta =  $\sum_{j=1}^M \sqrt{S_j P_j} / \sqrt{S_i P_i}$ ;
    object (i) .tag = object (i) .delta;
    object (i) .count =0;
    object (i) .enabled = TRUE;
    QueueObjectsByOrderIncreasingTag ();
end
FASE2 Multiplexación
while (not_exit)
    i =FindObjectWithLeastTagInActiveQueue ();
    SendObject(i);
    object (i) .tag = object (i) .tag + object (i) .delta;
    object (i) .enabled = FALSE;
    for j =1 to M
        Object (j) .count = object (j) .count +1;
        if (object (j) .count >= object (j) .delta)
            object (j) .count = object (j) .count - object (j) .delta;
            object (j) .enabled = TRUE;
        end
    end
end

```

fig. 5. Pseudocódigo del algoritmo Virtual Clock modificado

Para evaluar las prestaciones del algoritmo de multiplexación se han realizado simulaciones en las que la probabilidad de acceso a los ficheros por parte de los clientes es análoga a la que establece el servidor. En este caso, se cumplen las suposiciones establecidas en la obtención de (3) y se puede comparar las prestaciones del algoritmo de multiplexación de objetos con el límite teórico. En la siguiente gráfica se muestran los resultados obtenidos al evaluar el OAT con Virtual Clock. En la gráfica se muestran los resultados para  $OAT[FFQ, ZIPF(0.6), exponential]$  en rojo frente a  $OAT[MVT, ZIPF(0.6), exponential]$  en azul. Los resultados obtenidos con Virtual Clock modificado son en media un 26% superiores a los obtenidos con FFQ. También se han realizado simulaciones sin aplicar multiplexación, transmitiendo los objetos secuencialmente y evaluando el tiempo medio de acceso a los ficheros con una probabilidad de acceso uniforme. Los tiempos medios de acceso obtenidos son iguales a los obtenidos al evaluar (3):

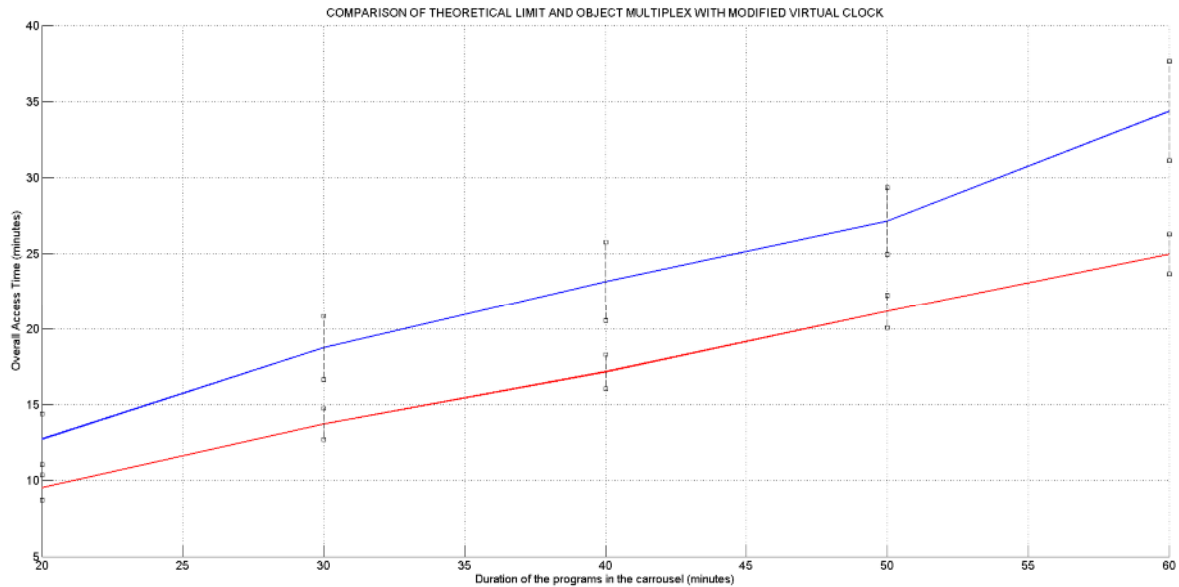


fig. 6. Comparación del Tiempo medio de Acceso obtenido al aplicar Virtual Clock con respecto al límite analítico.

En (3) se obtienen los tiempos medios de espera para todos los ficheros del carrusel. Al aplicar multiplexación de objetos, los ficheros más populares presentan un tiempo medio de espera menor que los menos populares. En las siguientes gráficas se muestran los tiempos de espera obtenidos en las simulaciones con  $OAT[MVC, ZIPF(0.6), constant]$  para los diez programas más populares en el carrusel, para diferentes valores de M. Se asume que la probabilidad de acceso local, por parte de los clientes, es igual para todos los objetos.

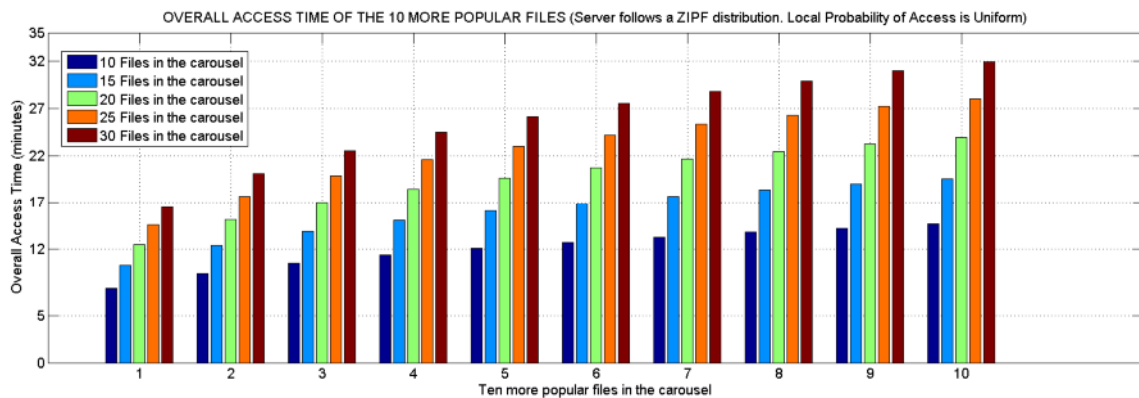


fig. 7. Tiempo medio de acceso a los 10 programas más populares para diferentes tamaños del carrusel

### V.3.EVALUACIÓN DE LOS MECANISMOS DE CACHING

En los resultados anteriores, el cliente recibe todos los ficheros del carrusel. Para modelar el efecto del cacheo de ficheros, el receptor debe seleccionar los ficheros del carrusel con una probabilidad diferente a la establecida por el transmisor. Para modelar la desviación en la probabilidad de acceso

a los ficheros por parte de los usuarios con respecto a la establecida por el servidor, se ha aplicado un método descrito en [3]. De acuerdo con este método, la probabilidad de acceso a los contenidos por parte de un usuario también sigue una distribución ZIPF. La diferencia entre las dos distribuciones ZIPF se modela a partir de dos parámetros. El primero, denominado *offset* representa la diferencia en la popularidad de dos ficheros, es decir, la diferencia del índice  $i$  en las distribuciones ZIPF para un fichero. Por ejemplo, el fichero más popular para el servidor ( $i=0$ ) es el sexto más popular ( $i=5$ ) para un usuario con *offset* = 5. El segundo parámetro, denominado *Noise* representa el porcentaje de archivos para los que las probabilidades de acceso son diferentes. Por ejemplo, para un valor de *Noise* del 50%, la mitad de los ficheros tendrán el mismo índice  $i$  en las distribuciones ZIPF, mientras que la otra mitad de los ficheros tendrán índices que se diferenciarán con un valor *offset*. El algoritmo de gestión de caché utilizado en las simulaciones es el *PIX*. Se asume que el recomendador es capaz de predecir con exactitud la probabilidad de acceso local, por lo que el valor de  $p$  para cada fichero se obtiene a partir de los valores de *Offset* y *Noise* del usuario. También se asume que el receptor conoce con exactitud el valor asignado para la tasa de transmisión de cada fichero. En la siguiente figura se muestra una comparación entre  $OAT[MVC, ZIPF(0.6), constant]$  (en rojo) y  $OAT[RR, ZIPF(0.6), constant]$  (en verde) para un tamaño de caché de 50MB y valores de *Offset* 5 y *Noise* 50% para las trazas en discontinua y *Offset* 0 y *Noise* 0% para las trazas en continuo. En azul se muestra  $OAT[FFQ, ZIPF(0.6), constant]$  sin caché:

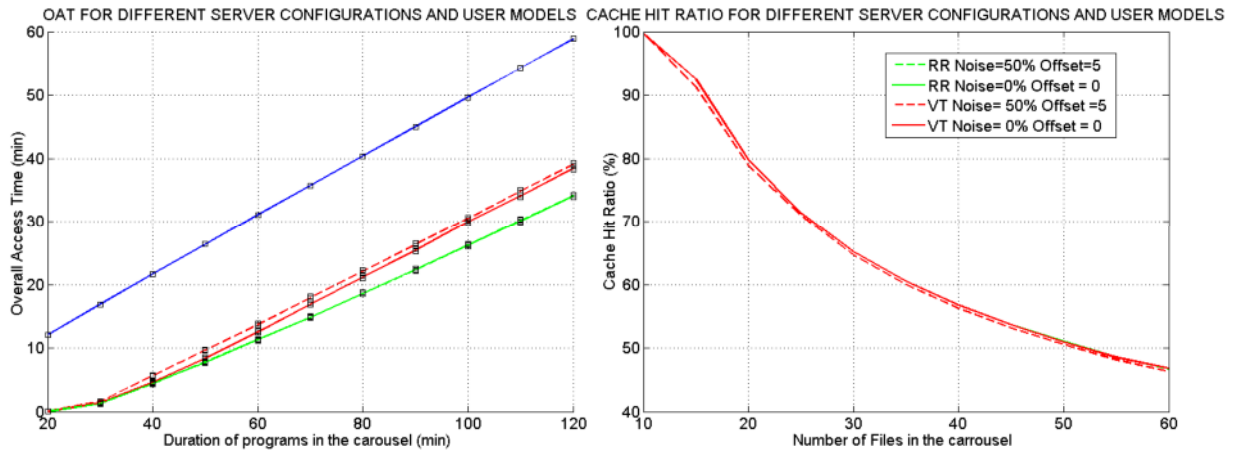


fig. 8. Comparación del tiempo medio de acceso con 50MB de caché entre multiplexación estadística y transmisión secuencial de los objetos

#### V.4.RESULTADOS DE LAS MEDIDAS EN LABORATORIO

A continuación, se presentan los resultados obtenidos de las medidas en laboratorio. En la siguiente gráfica se muestra el tiempo medio de acceso a los programas obtenido por medio de medidas, para un margen de confianza del 98%. La transmisión de los objetos se realiza de manera secuencial. El tamaño de los ficheros es una constante de 3.2MB que, de acuerdo con los métodos anteriores, se corresponde con un programa de aproximadamente 120 segundos de duración. Se han realizado

diferentes medidas para diferentes tamaños del carrusel, que se muestran en la siguiente gráfica junto con los valores esperados.

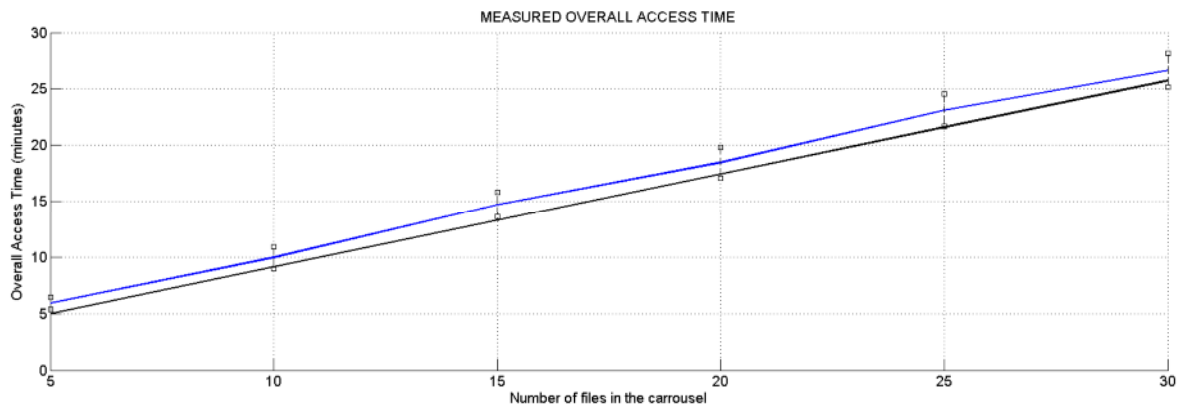


fig. 9. Tiempo medio de acceso medido en el laboratorio frente al esperado

## V.5. ANÁLISIS DE LOS RESULTADOS

**Escalabilidad con la oferta de contenidos:** Los diferentes resultados obtenidos demuestran que las mejoras introducidas mejoran la escalabilidad del servicio cuando aumenta la oferta de contenidos del carrusel. En la fig. 3 se muestra cómo la multiplexación de objetos proporciona mejoras en el tiempo medio de acceso con respecto a la transmisión secuencial. La figura también demuestra que la ganancia obtenida depende de la distribución de la popularidad de los contenidos y de sus tamaños. Los valores obtenidos muestran que la ganancia es mayor cuanto más dispares sean estos valores entre los diferentes programas a transmitir. Como se ha indicado, cabe esperar que la popularidad de los programas siga una distribución ZIPF y que los programas tengan duraciones diferentes, pero la configuración de estos parámetros no depende, en ningún caso, del operador del servicio, por lo que el margen de mejora que se introduce es circunstancial.

En la fig. 3 también se observa que las mejoras introducidas tienen carácter asintótico y que rondan valores entre el 70% y el 95% de los tiempos medios obtenidos a través de la transmisión secuencial. Como se muestra en la fig. 4 estas mejoras atenúan el efecto del tamaño del carrusel en los tiempos de espera, pero no modifican drásticamente el comportamiento del sistema, que no escala apropiadamente con el número de programas en el carrusel. Además, como muestra la fig. 6 la ganancia en el tiempo medio de acceso obtenida al aplicar multiplexación de objetos puede perderse íntegramente si el algoritmo utilizado para implementar la multiplexación no consigue ajustar con precisión las tasas establecidas. Este es el caso del algoritmo *Virtual Clock* modificado, que proporciona tiempos medios de espera similares al algoritmo *Round Robin* debido a que no consigue ajustar con precisión las tasas a largo plazo para la transmisión de los ficheros.

Es interesante observar el efecto de la multiplexación de objetos en presencia de una memoria caché en el cliente, gestionada a partir del algoritmo *PIX*. Como se puede apreciar en la fig. 8 los tiempos de acceso se reducen considerablemente cuando se utiliza una memoria en caché para cualquier disciplina en el servidor. La ganancia obtenida es proporcional al tamaño de la caché. Para las trazas con *Offset* 0 y *Noise* 0%, cuando los usuarios tienen una probabilidad de acceso a los contenidos similar a la establecida por el servidor, se observa que existe una diferencia entre los tiempos medios de acceso obtenidos por el algoritmo RR y el MVC y que estos últimos ofrecen peores resultados. Esta diferencia viene dada por la combinación de dos factores. Por una parte, los ciclos de transmisión de los contenidos menos populares son más largos y por tanto, el tiempo de acceso a ellos es mayor. Este hecho se muestra cuantitativamente en la fig. 7 para distintos tamaños del carrusel. Como se puede apreciar, aunque el tiempo medio de acceso a los contenidos sea equivalente al obtenido con la transmisión secuencial, el tiempo medio de acceso a los programas más populares es significativamente menor cuando se utiliza el algoritmo *Virtual Clock*, lo que se traduce en cierta ganancia en el tiempo de acceso a estos contenidos, en perjuicio de los menos populares, que tienen una tasa menor. Como se muestra en la fig. 8.b el número de aciertos de caché, es decir, el número de veces que un fichero que se solicita se encuentra íntegro en caché es equivalente para las dos disciplinas RR y MVC. Ahora bien, dado que MVC asigna menos tasa a los ficheros menos populares, éstos tienen más probabilidad de ser *cacheados* por un algoritmo basado en *PIX*. Esto no ocurre con RR, ya que todos los programas tienen asignados la misma tasa y el algoritmo de reemplazamiento en caché sólo tiene en cuenta la probabilidad de acceso a los programas.

Para MVC, si uno de los programas poco populares se encuentra en caché, la mejora en el tiempo de espera es considerable, pero el espacio que ocupa en caché previene que se almacenen otros ficheros más populares (que tengan un *PIX* menor) y el coste de un fallo de caché para un fichero con una tasa baja es más alto. En RR todos los objetos tienen las mismas tasas por lo que todos los fallos de caché tienen el mismo coste. La diferencia en el tiempo de espera entre las dos disciplinas con *PIX* dependerá de i) la distribución de la probabilidad de acceso a los programas e ii) la tasa asignada por la disciplina a cada uno de los programas, dado que ambos algoritmos obtienen tasas de acierto de caché equiparables. En el caso de una distribución ZIPF, los resultados que ofrece MVC son algo mejores que con RR, como se muestra en la fig. 8. Además, RR funciona mejor con valores de *Noise* y *Offset* altos.

Tras este análisis, se demuestra que el uso de MVC no mejora el tiempo medio de espera con respecto a RR en este caso de estudio.

**Usabilidad:** Existen dos modos de funcionamiento para un servicio de *TV/Cachecasting*, en el primero (Bajo Demanda) el usuario solicita qué programas quiere ver. Después, debe esperar cierto tiempo hasta que los contenidos estén disponibles en el terminal. En el segundo modo de funcionamiento, (Automático) el usuario accede directamente a la caché y consume el contenido



disponible en la memoria, de manera similar a la que se consumen programas de TV en los sistemas de difusión tradicionales.

Para el modo Bajo Demanda la usabilidad viene dada por el tiempo medio de acceso a los contenidos, dado que es el tiempo, en media, que el usuario debe esperar para acceder a un contenido. En base a los resultados obtenidos, se pueden discernir dos regiones en función del tiempo de espera que se considere como aceptable. En una de las regiones, los tiempos de espera son suficientemente bajos, aunque la oferta de contenidos en el carrusel es limitada. En la otra zona los tiempos medios de espera no se pueden considerar aceptables si consideramos que la aplicación debe funcionar estrictamente Bajo Demanda, pero por el contrario, la oferta de contenidos es más amplia. Si se entiende que los tiempos de espera en el entorno móvil deben ser muy pequeños, del orden de la duración de los programas, entonces la fig. 8 demuestra que el tamaño de la caché debe tener un tamaño similar al del carrusel de datos para que el porcentaje de acierto de caché ronde el 100%. En esa simulación, la memoria caché permite almacenar unos 13 ficheros en la caché. Cuando el tamaño del carrusel es de 15 ficheros, el tiempo de acceso es inferior a dos minutos mientras que para un tamaño de carrusel de 20 ficheros, el tiempo de acceso es de alrededor de 5 minutos.

Si el usuario accede directamente a los contenidos que hay en la memoria caché, la usabilidad de la aplicación viene dada por el tiempo transcurrido hasta que en la memoria caché hay una programación de contenidos que se ajusta en cierta medida a las preferencias del usuario. Conforme transcurra el tiempo, el algoritmo de gestión de caché irá almacenando los ficheros en el aire hasta ocupar el espacio en memoria. Una vez completado este proceso de carga inicial, el algoritmo de reemplazamiento de caché reemplazará programas, de manera que aumente la probabilidad de que los contenidos en la caché sean del agrado del usuario. Estos tiempos de carga estarán condicionados por el tiempo de acceso a los ficheros. Dado que con la multiplexación de objetos, los programas más populares tienen ciclos más cortos (fig. 7), y que el tiempo medio de acceso a los ficheros es del mismo orden de magnitud (fig. 8), los periodos de carga serán más cortos si se utiliza multiplexación de objetos. En opinión del autor, este modo debe ser el principal modo de funcionamiento para la aplicación.

## VI. CONCLUSIONES Y TRABAJO FUTURO

Los diferentes resultados generados han servido para evaluar las mejoras introducidas en los servicios de *TV/CacheCasting* por las técnicas de multiplexación de objetos y por los algoritmos de reemplazamiento en cache basados en *PIX*.

Los resultados obtenidos se pueden aplicar a dos casos: cuando no ocurren errores en la transmisión (poco probable en el entorno móvil) y cuando los clientes se conectan a un servicio de reparación de ficheros después de recibir el fichero por primera vez, para descargar al terminal los

fragmentos que no han sido recibidos [9]. Para minimizar la cantidad de ficheros que se deben descargar por la red de 3G como parte del proceso de reparación, se puede aplicar AL-FEC, tal y como se ha descrito en la sección IV.1 de este documento. Con pérdidas y AL-FEC, el tamaño de los objetos FLUTE que se utiliza en (3) se convierte en una variable aleatoria con cierta distribución de probabilidad en un área de servicio. Dado que AL-FEC añade símbolos de paridad por el carrusel de datos, la media de esta variable aleatoria será mayor que la que se ha aplicado en este trabajo, por lo que la configuración de AL-FEC permite balancear la carga entre el servidor FLUTE y el servidor de Reparación de Ficheros. El estudio de AL-FEC en la gestión de caché es el siguiente paso que debe seguir esta investigación.

En la metodología utilizada, se ha evaluado el tiempo medio de acceso a los ficheros. El modelo utilizado para caracterizar a los usuarios del servicio ha sido obtenido de la literatura ([3]) pero no se tienen ninguna justificación que motive el uso de este modelo, ni se han encontrado otros modelos que de manera justificada, modelen la probabilidad de acceso a los ficheros por parte de usuarios para este tipo de servicios. Se necesitan modelos más exactos para el comportamiento de los usuarios con respecto a los servicios de *TV/CacheCasting*. La red piloto de DVB-H desplegada en el campus de la Universidad Politécnica de Valencia (UPV) puede ser utilizada para realizar pruebas con usuarios finales y obtener estos modelos. Una vez que se cuenta con modelos más avanzados, se pueden analizar los tiempos de carga de caché descritos en V.5 y comprobar los tiempos transcurridos hasta que la memoria caché contiene una programación que se ajusta a las preferencias de los usuarios. Por último, se debe evaluar la influencia que tiene el diseño del recomendador que integra el gestor de caché en estos tiempos de carga. Como se ha mostrado en IV.1.1 y IV.1.2, la arquitectura del recomendador va a condicionar la predicción sobre la probabilidad de acceso futuro,  $P$ , que el recomendador debe proporcionar al gestor para implementar políticas de reemplazo en caché basadas en *PIX*. Con esto, se debe evaluar el efecto que tiene la arquitectura del recomendador en los tiempos de carga de la caché.

## **AGRADECIMIENTOS**

En esta sección, me gustaría expresar mi gratitud hacia todos mis compañeros de trabajo, sin los que no hubiera sido capaz de llevar a cabo este trabajo. En primer lugar, me gustaría agradecer a mi Director Juan Carlos Guerri por su devoción y su entusiasmo en la dirección de este trabajo, en concordancia con su labor como jefe del Grupo de Comunicaciones Multimedia (COMM). Después, debo agradecer a los miembros de dicho grupo Ismael, Pau, Patricia, Román, Tito, Victor y Wilder (en orden alfabético) por su compañerismo. En especial, quisiera agradecer a Ismael sus correcciones, su interés y su preocupación por mi tesina. A Román por su gran ayuda con las medidas de laboratorio y a Pau por sus consejos para la presentación de los resultados.

Debo expresar mi gratitud al Ministerio de Industria, Turismo y Comercio por su apoyo al proyecto ModelaTV, en el que se enmarca este trabajo de investigación.

También quería expresar mi gratitud a mi familia y mis amigos, a mis padres Manuel y Pascuala y a mi hermana Celia, por su apoyo y paciencia.

## REFERENCIAS

- [1] G. Zhiqi, Y. Songyu, Z. Weinjun *Using Object Multiplex Technique in Data Broadcast on Digital CATV Channel*. IEEE Transactions on Broadcasting, vol. 50. No. 2, June 2004.
- [2] E. A. Heredia, *Optimal Object Allocation for Multimedia Broadcast*, IEEE International Conference on Speech, Acoustics, and Signal Processing, 1998.
- [3] S. Acharya, R. Alonso, M. Franklin, S. Zdonik, *Broadcast Disks: Data Management for Asymmetric Communication Environment*, ACM SIGMOD Conference, San Jose, Estados Unidos, Mayo 1995.
- [4] ETSI EN 302 304 v1.1.1 *DVB-H - Transmission System for Handheld Terminals*, Noviembre 2004.
- [5] ISO/IEC 13818-1, *Information Technology — Generic Coding of Moving Pictures and Associated Information: Systems*, Diciembre 2000
- [6] ETSI EN 300 468 V1.9.1, *Digital Video Broadcasting (DVB); Specification for Service Information (SI) in DVB systems*, Marzo 2009
- [7] ETSI EN 301 192 v1.4.2, *Digital Video Broadcasting (DVB); DVB specification for data broadcasting*, Abril 2008
- [8] ETSI TS 102 472 V1.2.1, *Digital Video Broadcasting (DVB); IP Datacast over DVB-H: Content Delivery Protocols (CDP)*, Diciembre 2006
- [9] OMA-TS-BCAST\_Distribution-V1\_020090212-A, *File and Stream Distribution for Mobile Broadcast Services*, Febrero 2009
- [10] IETF RFC 3926, *FLUTE - File Delivery over Unidirectional Transport*, Octubre 2004.
- [11] OMA-TS-BCAST\_Service\_Guide-V1\_0-20090212-A, *Service Guide for Mobile Broadcast Services*, Febrero 2009.
- [12] ETSI TS 102 471 V1.2.1 *Digital Video Broadcasting (DVB); IP Datacast over DVB-H: Electronic Service Guide (ESG)*, Noviembre 2006.
- [13] Paila, T., *Unidirectional IP-based mass file delivery protocol*, Wireless Communication Systems, 2004. 1st International Symposium on, vol., no., pp. 235-239, 20-22 Sept. 2004
- [14] IETF RFC 3450, *Application Layer Control (ALC) Protocol Instantiation*, Diciembre 2002
- [15] IETF RFC 3048, *Reliable Multicast Transport Building Blocks for One-to-Many Bulk-Data Transfer*, Enero 2001.
- [16] IETF RFC 3451, *Layered Coding Transport (LCT) Building Block*, Diciembre 2002.
- [17] IETF RFC 3452, *Forward Error Correction (FEC) Building block*, Diciembre 2002.
- [18] IETF RFC 2327, *Session Description Protocol*, Abril 1998.
- [19] IETF RFC 5053, *Raptor Forward Error Correction Scheme for Object Delivery*, Octubre 2007
- [20] G. Adomavicious, A. Tuzhilin, *Towards the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions*, IEEE Transactions on Knowledge and Data Engineering, vol. 17, No. 6, Junio 2005.
- [21] ETSI TS 102 822-2 V1.4.1, *Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 2: Phase 1 - System description*, Noviembre 2007

- [22] G. Salton, *Automatic Text Processing*, Addison-Wesley, 1989.
- [23] G. K. Zipf, *Selected Studies of the Principle of Relative Frequency in Language*. Cambridge, Estados Unidos, 1932.
- [24] M. Chesire, A. Wolman, G. M. Voelker, H. M. Levy, *Measurement and Analysis of a Streaming-Media Workload*, 3<sup>rd</sup> Usenix Symposium on Internet Technologies and Systems, March 2001.
- [25] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, *Web Caching and Zipf-like Distributions: Evidence and Implications*, proceedings of IEEE INFOCOM , Marzo 1999
- [26] C. Södergård, *Mobile Television – Technology and User Experiences, Report on the Mobile-TV Project*, VTT Publications, 2003.

## ANEXOS

El artículo que se incluye como anexo se presenta en la séptima edición del congreso internacional *European Interactive TV Conference* (EuroITV 2009) en Junio de 2009. El artículo se enmarca en la línea de *Systems and Enabling Technologies*.

- [1] F. Fraile, I. de Fez, J.C. Guerri *ModelaTV – Service Personalization and Business Model Management for Mobile TV*, 7<sup>th</sup> European Interactive TV Conference, Leuve Bélgica, Junio 2009.

Además, el autor ha contribuido en la integración del servicio de transferencia de ficheros en la gestión de emergencias, que se muestra en el siguiente artículo, presentado en el 34<sup>th</sup> EUROMICRO Conference on Software Engineering and Advanced Applications (SEAA) en 2008:

- [2] R. Belda, I. de Fez, F. Fraile, V. Murcia, P. Arce, J. C. Guerri, *Multimedia System for Emergency Services over TETRA-DVBT Networks* 34<sup>th</sup> EUROMICRO Conference on Software Engineering and Advanced Applications (SEAA) Parma, Italy, September 2008