Estudio y Optimización de los Procedimientos de Adaptación al Enlace en HSDPA

Autor: Martín-Sacristán Gandía, David

Director: Cardona Marcet, Narcís

Resumen — La tecnología HSDPA (High Speed Downlink Packet Access) es una evolución de UMTS creada con el objetivo de aumentar la capacidad de transmisión en el enlace descendente. Su mejora se basa en la utilización de un canal compartido de comunicación gestionado de forma eficiente desde la estación base (por medio de un *packet scheduler*), la utilización de mecanismos de retransmisión y combinación de información avanzados (*hybrid ARQ*) y la posibilidad de emplear modulaciones de alto orden (16QAM y 64QAM).

Las dos últimas características nombradas serían inútiles sin unos buenos procedimientos de adaptación al enlace (*link adaptation*) que ajustaran los parámetros de transmisión a la calidad del enlace radio.

La presente tesina aborda el estudio y optimización de los mecanismos de *link adaptation* en HSDPA. Para tratar el problema se siguen dos estrategias. Por un lado, se estudia un *link adaptation* genérico con el fin de obtener conclusiones fácilmente trasladables a sistemas particulares como HSDPA. Por otro lado, se aportan soluciones a problemas específicos de HSDPA como los fallos del *link adaptation* con baja carga.

Abstract — HSDPA (High Speed Downlink Packet Access) technology is an evolved version of UMTS focused on the improvement of the downlink capacity. HSDPA enhancement is based on the efficient management of a shared channel done by the Node-B (employing a packet scheduler), the using of advanced retransmission and combination mechanisms (hybrid ARQ) and the availability of high order modulations (16QAM and 64QAM).

The later characteristics would be worthless without good link adaptation procedures that adjust transmission parameters according to the radiolink quality.

This thesis deals with the study and optimization of link adaptation mechanisms in HSDPA. Two strategies are followed herein. First, a generic link adaptation is studied with the aim of reaching some general conclusions and applying them to real systems as HSDPA. Besides, a more detailed study is done for HSDPA finding solutions for some specific problems as link adaptation failures with low load.

Autor: Martín-Sacristán Gandía, David; email: damargan@iteam.upv.es
Director: Cardona Marcet, Narcís; email: ncardona@dcom.upv.es
Fecha de entrega: 3-12-07

I. INTRODUCCIÓN.

El mercado actual no se contenta con obtener de las redes de comunicaciones móviles un servicio básico de comunicación en movilidad. Más bien, los usuarios esperan recibir un servicio equiparable al de las redes fijas. Esta demanda fuerza y orienta la evolución de las diferentes tecnologías. En este sentido, el estándar 3G UMTS ha evolucionado hacia el HSDPA (*High Speed Downlink Packet Access*) con el objetivo de aumentar la tasa de transmisión en el enlace descendente.

Si bien muchos de los procesos y algoritmos de HSDPA están claramente especificados, para garantizar la interoperabilidad entre equipos y sistemas, algunos puntos clave están abiertos a diferentes implementaciones. Este hecho ha propiciado el desarrollo de una gran actividad investigadora en campos como el *link adaptation* (adaptación al enlace) o el *scheduling* (planificación de las transmisiones).

En esta tesina se aborda el estudio y mejora de los mecanismos de *link adaptation* en HSDPA.

En la sección II se aclararán ciertos conceptos que darán al lector una visión global de la materia bajo estudio. Concretamente, se hará una breve descripción de HSDPA y se definirán los principales conceptos relacionados con la gestión de recursos radio en redes celulares.

En la sección III será la primera dedicada totalmente al *link adaptation* y en ella se realizará una exposición de su estado del arte. Para ello se presentarán las técnicas de *link adaptation* más comunes adoptadas en GPRS, EDGE y HSDPA.

A continuación, en la sección IV, se modelará y analizará un *link adaptation* genérico. El objetivo de este análisis será la obtención de una serie de conclusiones fácilmente aplicables a sistemas reales como, por ejemplo, HSDPA.

La sección V presentará los aspectos más relevantes del entorno de simulación empleado. A este respecto, es de destacar que el presente estudio se ha desarrollado como la continuación lógica del proyecto final de carrera del autor de la tesina, titulado: ‘Diseño de un simulador de sistema para el enlace descendente de HSDPA’. Durante la realización de la tesina se han introducido cambios en el simulador ya desarrollado, lo cual ha permitido una evaluación realista de las soluciones aportadas al *link adaptation* en HSDPA.

En la sección VI se estudiará en profundidad el *link adaptation* en HSDPA. Se discutirán las ventajas e inconvenientes del esquema típico y se propondrán una serie de mejoras. Para obtener tales mejoras se aplicarán tanto las conclusiones extraídas del análisis genérico previo como nuevas soluciones adoptadas tras el estudio particular de HSDPA. Todas las estrategias propuestas serán comparadas en base a los resultados de simulaciones.

Por último, en la sección VII, se resumirá el estudio realizado y se presentarán las conclusiones más relevantes.

II. CONCEPTOS BÁSICOS.

II.1. *High Speed Downlink Packet Access (HSDPA).*

En este apartado se destacan los aspectos de HSDPA relevantes para entender el presente estudio. Una buena descripción general de HSDPA se puede encontrar en [1] y [2].

En primer lugar es de destacar que HSDPA emplea un canal de transporte compartido. No se reservan recursos estableciendo un canal dedicado a un usuario, sino que el conjunto de los recursos del sistema se reparte dinámicamente entre los usuarios. Cada 2 ms, que es el tiempo de transmisión de un bloque radio (TTI o *Transmission Time Interval*), se reasignan códigos de espectro ensanchado entre los usuarios del sistema. Se permite tanto la asignación de múltiples códigos a un usuario (transmisión multicódigo) como la asignación simultánea de códigos a varios usuarios (multiplexación por división en código). Esta reasignación rápida se conoce comúnmente como *fast scheduling* y es claramente beneficiosa para la transmisión de paquetes.

En HSDPA es posible adaptar la modulación y codificación dinámicamente, en cada TTI. Las modulaciones permitidas son QPSK, 16QAM y en las últimas versiones del estándar 64QAM. La disponibilidad de modulaciones de alto orden junto con el amplio ancho de banda empleado (alrededor de 5MHz) permite altas velocidades de transmisión. La codificación básica se obtiene con un turbocodificador de tasa 1/3, pero es posible obtener otras tasas empleando repetición o perforado (*puncturing*) de bits. Se deduce que hay una gran cantidad de combinaciones de modulación y tasa de codificación. Por un lado esto permite una gran adaptabilidad al enlace radio y dota al sistema de una alta flexibilidad para dar servicio a usuarios con diferentes requerimientos de servicio.

Una novedad de HSDPA es el uso del mecanismo *Hybrid ARQ*. Este mecanismo incluye un ARQ (*Automatic Repeat-reQuest*) tradicional en el que el receptor envía reconocimientos al emisor para informarle de la recepción correcta o incorrecta de los datos, requiriendo o no una retransmisión. Concretamente, esta tarea se realiza por medio de un protocolo de tipo *stop-and-wait* (SAW) multicanal en el que varios procesos SAW operan en paralelo. Cada proceso SAW en el emisor realiza paradas tras cada transmisión hasta la recepción de su reconocimiento, de ahí la necesidad de los múltiples canales para realizar un uso eficiente de los recursos. Pero el mecanismo *Hybrid ARQ* permite además que las retransmisiones de una misma información se combinen en el receptor aumentando la probabilidad de realizar una decodificación correcta. La combinación puede darse según dos esquemas: *Chase combining*, en el que las transmisiones a combinar son idénticas, e *Incremental Redundancy*, en el que las transmisiones a combinar son diferentes.

II.2. *Gestión de recursos y control de enlace en redes celulares.*

En esta sección se introducen ideas básicas sobre gestión de recursos radio y sobre control del enlace en un sistema celular.

La gestión de recursos radio (Radio Resource Management – RRM) comprende un conjunto de algoritmos y procesos que el sistema celular emplea para repartir los recursos radio disponibles entre los usuarios del mismo. En el caso concreto de HSDPA, los recursos que deben ser gestionados en la estación base son: códigos, potencia y tiempo. En cuanto a los códigos, se dispone de un máximo de 15 códigos de espectro ensanchado con factor de ensanchado 16 los cuales se pueden distribuir entre los usuarios del sistema en cada periodo de asignación de recursos. Igualmente hay una potencia disponible para transmisión de datos en la celda, cuya disponibilidad puede fluctuar o no con el paso del tiempo (hay canales comunes consumiendo también potencia e incluso comunicaciones por canales dedicados de UMTS). En cuanto al tiempo, se puede considerar como un recurso más del sistema que se otorga con una granularidad de 2 ms (periodo de asignación de recursos).

De forma genérica, en un sistema de transmisión de paquetes, el *packet scheduler* decide qué paquetes de entre los que están disponibles para transmitirse deben enviarse en cada momento. En una red celular suele identificarse este concepto con el conjunto de algoritmos y técnicas ejecutados de forma periódica para repartir los recursos del sistema entre los usuarios de transmisión de paquetes [3]. Concretamente en HSDPA, el *packet scheduler* es el principal gestor de los recursos del Nodo-B dedicados a la transmisión de información de usuario. El *scheduler* decide qué usuarios del sistema serán destinatarios de las transmisiones (con lo que consumirán los recursos del mismo) en cada intervalo TTI y qué cantidad de códigos y potencia se le asignará a cada usuario. Atendiendo al número de usuarios a los que se da recursos en cada periodo de asignación de recursos, se podrían clasificar las estrategias de *scheduling* en TDM o *Time Division Multiplex*, en las que los recursos se otorgan únicamente a un usuario en cada periodo de asignación, y CDM o *Code Division Multiplex*, en las que en cada periodo de asignación se otorgan recursos a varios usuarios multiplexados en código. Un buen resumen sobre algoritmos de *scheduling* en HSDPA se puede encontrar en [4].

La adaptación al enlace, comúnmente denominada *link adaptation* (LA), consiste en todos aquellos mecanismos que permiten modificar dinámicamente los parámetros de las transmisiones individuales (entre el emisor y un receptor) realizadas en el sistema con el objetivo de optimizar las prestaciones de dichas transmisiones de acuerdo a algún criterio. Un criterio válido podría ser, por ejemplo, la maximización del *throughput* de la transmisión, entendido como número de bits correctamente recibidos por unidad de tiempo. Algunos de los mecanismos que se podrían clasificar como de *link adaptation* son los siguientes:

- *Adaptive Modulation and Coding* (AMC): cambio de modulación y codificación.

- *Incremental Redundancy* (IR): cambio en la redundancia de las transmisiones
- *Adaptive Power Allocation* (APA) y *Power Control* (PC): gestión dinámica de la potencia.

Si bien el concepto de *link adaptation* ha sido bastante tratado en la literatura la idea que se tiene de él es muy diferente según el autor y el sistema que trate. En la presente tesina se estudiará el *link adaptation* desde la perspectiva del AMC e IR dejando aparcado el APA por ser menos relevante en HSDPA.

III. ESTADO DEL ARTE EN LINK ADAPTATION.

En la sección anterior se definió el *link adaptation* de forma genérica. En la presente sección se describirán algunos mecanismos de *link adaptation* reales de sistemas actualmente desplegados.

III.1. *GPRS.*

La tecnología GPRS (General Packet Radio Service) surgió como una evolución de la red GSM permitiendo la transmisión eficiente de comunicaciones de paquetes. En GPRS se pueden emplear cuatro esquemas de codificación (*Coding Schemes – CS*), todos ellos con la misma modulación pero diferente tasa de codificación. Se recogen en la siguiente tabla extraída de [5]:

Esquema	Modulación	Tasa de codificación	Bits de datos por bloque radio	Tasa (kbps)
CS1	GMSK	1/2	181	9.05
CS2	GMSK	~ 2/3	268	13.4
CS3	GMSK	~ 3/4	312	15.6
CS4	GMSK	1	428	21.4

Tabla 1: Esquemas de codificación en GPRS

El *link adaptation*, en este caso, realiza un AMC que trata de elegir el esquema de codificación más adecuado a la calidad radio, generalmente con el objetivo de maximizar el *throughput*. Para ello se suelen emplear las curvas de *throughput* frente a CIR (*Carrier to Interference power Ratio)* de cada modo. Es importante tener en cuenta que las curvas de *throughput* dependen de la velocidad del móvil y de la aplicación o no de *frequency hopping* entre otros factores. La Fig.1 muestra estas curvas para un usuario a 3 km/h y con *frequency hopping.*

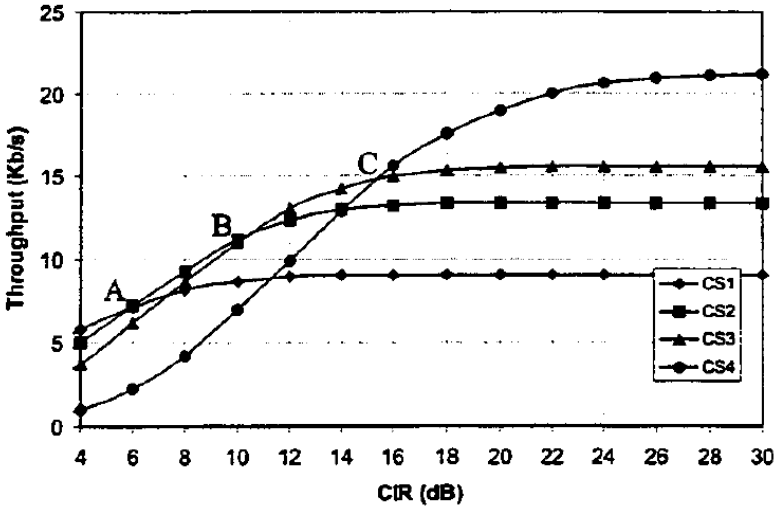

Fig.1. Curvas de throughput frente a CIR para los esquemas de codificación de GPRS [6].

Hay dos tipos de algoritmos clásicos estudiados para el *link adaptation* en GPRS:

- Algoritmos basados en CIR (CIR-*based*): se emplea el valor de CIR como indicador de la calidad del enlace. Se definen unos umbrales de CIR correspondientes a las intersecciones de las curvas de *throughput* para los diferentes CS.

  Este método es equivalente a elegir el CS que maximice el *throughgput* efectivo. Este *throughput* se puede calcular con la siguiente fórmula [6]:

$$Thr = Rate(CS) * \left(1 - BLER(CS, CIR)\right) \tag{1}$$

- Algoritmos basados en BLER (BLER-*based*): el parámetro empleado para estimar la calidad del canal es la proporción de bloques erróneamente recibidos, o BLER (*Block Error Rate*), promediada durante un cierto periodo de tiempo. Se puede encontrar un ejemplo en [6]. En dicho artículo se presenta una fórmula para obtener los umbrales de BLER de un modo *n*:

$$BLER_{n,\max} = \left(1 - \frac{Thr_{n,\inf}}{Thr_{n,\max}}\right) \tag{2}$$

$$BLER_{n,\min} = \left(1 - \frac{Thr_{n,\sup}}{Thr_{n,\max}}\right) \tag{3}$$

  ,donde $BLER_{n,max}$ y $BLER_{n,min}$ son los umbrales de BLER máximo y mínimo para el modo *n*, $Thr_{n,max}$ es la tasa máxima del modo *n*, y $Thr_{n,inf}$ y $Thr_{n,sup}$ son los *throughputs* del modo *n* en los puntos de intersección con las curvas de *throughput* de los modos inferior y superior. El algoritmo funciona de modo que si el BLER estimado durante un periodo de referencia es menor que el mínimo se elige un modo más eficiente (el modo superior), si el BLER es mayor que el máximo se elige un modo más robusto (el modo inferior) y en otro caso se mantiene el modo en uso.

En general, las prestaciones del algoritmo CIR-based son superiores a las del BLER-based para velocidades altas y similares para velocidades bajas. La definición de unos buenos umbrales es clave, pero dichos umbrales varían según se considere o no frequency-hopping y según la velocidad, lo cual complica bastante la optimización del *link adaptation.*

Información más detallada sobre el *link adaptation* en GPRS se puede encontrar en [7].

En la práctica, como medida de calidad para realizar el *link adaptation*, se ha empleado la medida denominada RXQUAL. Este parámetro es una medida del *pseudo bit error rate* (PBER), que se calcula comparando la secuencia recibida (antes de decodificar) con la secuencia decodificada y vuelta a codificar. Contando el número de errores (diferencias) se obtiene dicho PBER.

III.2. *EDGE.*

En EDGE (Enhance Data Rates for GSM Evolution) se pueden emplear hasta 9 esquemas de modulación y codificación (*Modulation and Coding Scheme* o MCS) con las siguientes características obtenidas de [5]:

Esquema	Modulación	Tasa de codificación	Familia	Tasa (kbps)
MCS1	GMSK	0.53	C	8.8
MCS2	GMSK	0.66	B	11.2
MCS3	GMSK	0.85	A padding A	13.6 14.8
MCS4	GMSK	1.00	C	17.6
MCS5	8PSK	0.37	B	22.4
MCS6	8PSK	0.49	A padding A	27.2 29.6
MCS7	8PSK	0.76	B	44.8
MCS8	8PSK	0.92	A padding	54.4
MCS9	8PSK	1.00	A	59.2

Tabla 2: Esquemas de modulación y codificación en EDGE

Se pueden obtener curvas de *throughput* frente a SIR similares a las de GPRS para cada MCS. Pero en este caso la peculiaridad más destacable de EDGE es que permite combinar las diferentes retransmisiones de los paquetes recibidas mediante *incremental redundancy* (IR).

En EDGE se definió una nueva medida de calidad diferente al RXQUAL previo: el BEP (*Bit Error Probability*), más útil para medir la calidad en transmisión de paquetes. Esta medida se evalúa en cada ráfaga, obteniéndose mayor información sobre el canal radio y su variación durante la transmisión de un bloque de datos. Empleando el BEP de cada ráfaga (BEP_{Burst}) se obtienen el $MEAN\_BEP_{block}$ (media) y CV_BEP_{block} (varianza) de cada bloque radio tal y como se indica en

[8]. Promediando dichos valores durante un periodo de reporte se obtienen el MEAN_BEP y CV_BEP. En el anexo D de [8] se aporta un ejemplo de selección de esquema de modulación y codificación basado en las estimas MEAN_BEP y CV_BEP orientado a la maximización del *throughput*. Es necesario disponer de tablas de este tipo con y sin IR, y para cada tipo de modulación.

AMC e IR son complementarios en EDGE como se estudia en [9]. En dicho artículo se explica que un IR puro que emplee modulaciones de orden superior puede obtener mayores tasas que un LA puro con modulaciones de diferentes órdenes, pero a costa de una mayor complejidad (más memoria) y pudiendo presentar un aumento en el retardo de la transmisión. Tanto el retardo como los requisitos de memoria pueden reducirse empleando tasas de codificación iniciales bajas, pero de esta forma el *throughput* disminuiría. Además de los problemas reseñados, por cuestiones de robustez es necesario tener la posibilidad de emplear modulaciones de bajo orden si la calidad del canal es muy baja, para asegurar la recepción de la información. De esta exposición se deduce que el *link adaptation* es necesario pero que debe ser optimizado.

III.4. *HSDPA.*

En esta subsección se presenta el marco de referencia en el que se desarrolla el *link adaptation* en HSDPA y el algoritmo empleado comúnmente por los investigadores de Nokia.

III.4.1 *Channel Quality Indicator*

Como se ha visto para otros sistemas, todo algoritmo de *link adaptation* emplea algún indicador de la calidad del canal como parámetro de entrada. El mecanismo de adaptación al enlace utilizado en HSDPA se basa en información de *feed-back* que los terminales móviles mandan al Nodo-B. La diferencia con respecto a otros sistemas radica en que esta información no consiste en un valor de CIR o SIR (*Signal to Interference power* Ratio) o similar medido por el terminal. La filosofía es distinta en HSDPA. En su lugar, el terminal recomienda al Nodo-B una combinación de parámetros de transmisión empleando un indicador conocido como *Channel Quality Indicator* (CQI). En las especificaciones de HSDPA [10] se puede encontrar la definición, procedimiento de medida y reporte del CQI. A continuación se incluyen los conceptos más relevantes para el presente estudio.

El CQI es un valor entero entre 0 y 30 que se envía desde los terminales móviles al Nodo-B. Físicamente está contenido en dos slots del canal ascendente llamado HS-DPCCH (High Speed Downlink Physical Common Control Channel), concretamente en los dos últimos slots de cada subtrama radio (formada por tres slots).

Cada CQI tiene una traducción directa a una combinación de modulación, número de códigos, tamaño de bloque de transporte y tasa de código recogida en tablas contenidas en [10], que en adelante se denominará ‘modo asociado’ al CQI.

La generación del CQI es de la siguiente forma. El terminal móvil realiza medidas de nivel físico sobre el canal piloto de UMTS (CPICH) durante un período que acaba 1 slot antes de empezar el reporte del CQI, esto es, durante la subtrama anterior a aquella en la cual se produce el reporte. Con las medidas de nivel físico se obtiene la SINR del canal piloto. Dicha SINR se traduce a un valor de SINR de HSDPA mediante la adición de una corrección de potencia que se asume es la diferencia de potencia entre el canal piloto y la potencia de una transmisión de datos. Este valor viene configurado por capas superiores. Además se especifica que el terminal debe asumir que la potencia estaría igualmente distribuida entre todos los códigos asignados en una transmisión. Con estos datos, cada terminal debe ser capaz de obtener el CQI cuyo modo asociado asegurara que, para la calidad de canal experimentada durante el período de medida, se tuviera un valor de BLER menor al 10% en recepción. El método que el terminal siga para cumplir con este objetivo no está especificado, lo que sí está especificado son unos requisitos mínimos que se deben cumplir bajo ciertas condiciones de test referentes a la relación entre el CQI reportado y el BLER experimentado, recogidos en [11].

Dado que el cálculo del CQI se apoya en la estimación de una SINR es importante tener en cuenta que la estimación de la SINR en la realidad no es perfecta existiendo un error en la medida.

Otro aspecto de los CQIs que será relevante en adelante es todo lo relativo a su gestión de tiempos. En primer lugar, ya se ha dicho que la medida de canal se realiza durante 2ms. El informe del CQI finaliza 2ms más tarde. Además, hay un retardo adicional en el Nodo-B desde la recepción del CQI hasta que se efectúa una transmisión teniéndolo en cuenta. En conjunto, se asume que el tiempo que pasa desde la medida de un CQI hasta que se recibe en el terminal móvil una transmisión formateada en base a la información de ese CQI es de 7 ms. Este retardo, considerando que el canal móvil es variable puede ser muy perjudicial. Es destacable que en ese cálculo se asume que el CQI se envía en todas las subtramas ascendentes, esto es, con una periodicidad de 2ms. Si bien esta configuración es común en las redes actualmente desplegadas otros valores son posibles. La Fig. 2 recoge el diagrama temporal del cálculo y envío del CQI tal y como se ha explicado.

En cuanto a los modos asociados a los CQIs, existen diferentes tablas que los relacionan según la categoría del terminal HSDPA. Estas tablas se diseñaron con una propiedad común consistente en que la diferencia entre los valores de SINR que precisan modos consecutivos para un BLER del 10% es de aproximadamente 1dB. En la Fig.3 se presentan las curvas de BLER frente a SINR para los 30 modos de la categoría 10 de Terminal, donde se observa fácilmente la diferencia de 1dB.

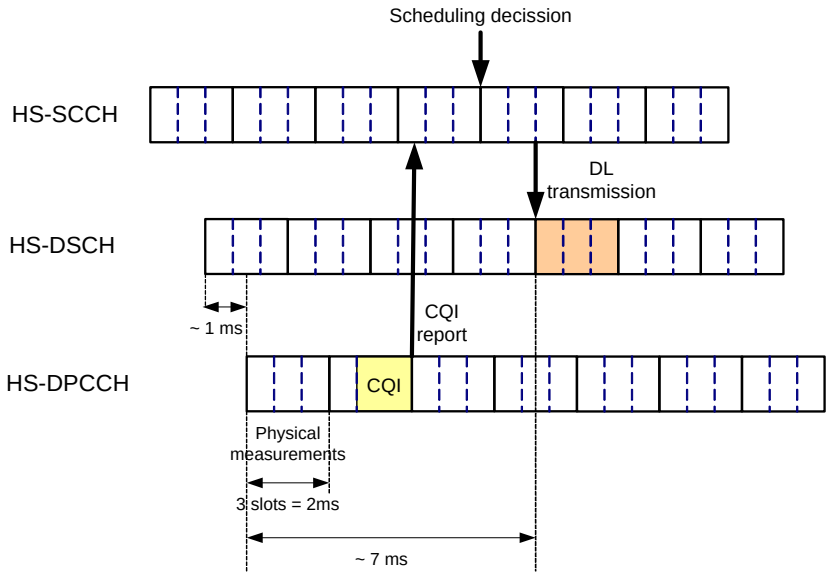

Fig.2. Diagrama temporal en el procesado de los CQIs.

CQI	TBS	Número de códigos	Modulación	Tasa (kbps)	Tasa de codificacion
1	137	1	QPSK	68.5	0.168
2	173	1	QPSK	86.5	0.205
3	233	1	QPSK	116.5	0.268
4	317	1	QPSK	158.5	0.355
5	377	1	QPSK	188.5	0.418
6	461	1	QPSK	230.5	0.505
7	650	2	QPSK	325.0	0.351
8	792	2	QPSK	396.0	0.425
9	931	2	QPSK	465.5	0.497
10	1262	3	QPSK	631.0	0.447
11	1483	3	QPSK	741.5	0.523
12	1742	3	QPSK	871.0	0.613
13	2279	4	QPSK	1139.5	0.600
14	2583	4	QPSK	1291.5	0.679
15	3319	5	QPSK	1659.5	0.696
16	3565	5	16-QAM	1782.5	0.374
17	4189	5	16-QAM	2095.5	0.439
18	4664	5	16-QAM	2332.0	0.488
19	5287	5	16-QAM	2643.5	0.553
20	5887	5	16-QAM	2944.5	0.616
21	6554	5	16-QAM	3277.0	0.685
22	7168	5	16-QAM	3584.0	0.749
23	9719	7	16-QAM	4859.5	0.725
24	11418	8	16-QAM	5709.0	0.745
25	14411	10	16-QAM	7206.5	0.752
26	17237	12	16-QAM	8774.0	0.763
27	21754	15	16-QAM	10877.0	0.756
28	23370	15	16-QAM	11685.0	0.812
29	24222	15	16-QAM	12111.0	0.842
30	25558	15	16-QAM	12779.0	0.888

Tabla 3: Tabla de CQI para la categoría 10 ([10])

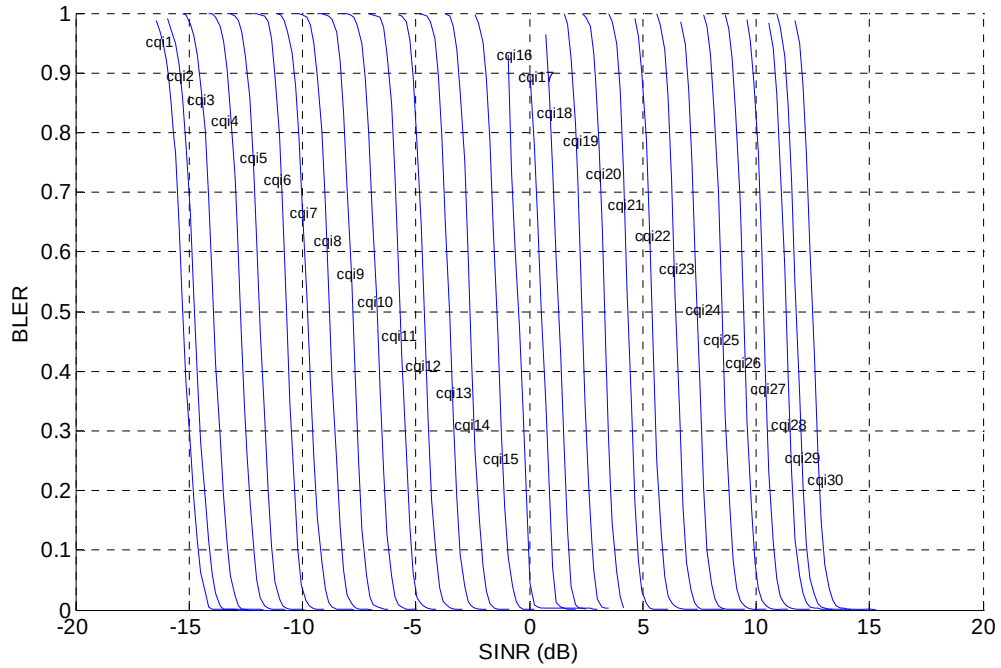

Fig.3. Curvas de throughput frente a SINR para los modos indicados en la Tabla 3.

III.4.1 *Algoritmo de Link Adaptation*

El papel del CQI enviado por el terminal móvil al Nodo-B es recomendar una combinación de parámetros de transmisión. Sin embargo, esta recomendación no es vinculante para el algoritmo de *link adaptation* que puede tomar decisiones en base a otros parámetros adicionales.

El algoritmo de LA descrito en [12] utiliza la información contenida en los CQIs enviados por los terminales móviles al Nodo-B pero hace uso de un bucle de control externo que corrige los CQIs reportados para asegurar un BLER objetivo en la transmisión. Otra descripción del mismo puede encontrarse en [1]. El algoritmo estudiado está en consonancia con lo propuesto por otros autores en [13] y [14].

La estrategia pasa por tener un factor de corrección A, expresado en decibelios, que se va actualizando con los reconocimientos del protocolo ARQ. Al recibirse un ACK, A se reduce en A_{down} decibelios y al recibirse un NACK, A se incrementa en A_{up} decibelios.

Al recibirse un CQI éste se traduce en un valor de SIR en base a tablas que contiene el Nodo-B. A dicho valor de SIR se le resta el valor de A, obteniendo un SIR corregido. El CQI finalmente considerado será aquel cuyo valor de SIR correspondiente se acerque más al corregido. Para que esta estrategia sea realizable se debe asumir un conocimiento a priori de la relación entre CQI y SIR en el Nodo-B. En [12], en lugar de asumir el conocimiento exacto de esta relación se asume el conocimiento de la diferencia relativa de SIR entre CQIs consecutivos fijándola a 1dB (consecuente con lo que se vio en el apartado anterior), lo cual viene a ser equivalente.

Con el algoritmo propuesto, en caso de que la estimación del CQI enviada por el móvil al Nodo-B sea sistemática o estadísticamente optimista, el factor de corrección tenderá a crecer al producirse

errores en la transmisión y recibirse NACKs, lo que hará que el CQI corregido tienda a ser menor que el indicado por el móvil, con lo que la probabilidad de que se produzcan errores en la transmisión decrecerá.

Se puede demostrar (aunque se omite por razones de espacio) que usando este algoritmo el BLER de las primeras transmisiones (sin tener en cuenta el BLER de las retransmisiones) tiende a:

$$BLER = \frac{1}{1+\frac{A_{up}}{A_{down}}} \tag{4}$$

IV. ESTUDIO DE UN LINK ADAPTATION GENÉRICO.

IV.1 *Presentación del método.*

A continuación se detalla un procedimiento genérico de *link adaptation* basado en AMC y diseñado con el objetivo de maximizar el *throughput* efectivo de la transmisión. Este LA genérico es muy similar al empleado en GPRS. Los conceptos estudiados y mecanismos de mejoras serán fácilmente trasladables a otros sistemas como HSDPA.

Como se ha indicado, el LA genérico bajo estudio estaría basado en AMC, es decir, habría diferentes modos de transmisión cada uno caracterizado por una diferente combinación de modulación y codificación. El procedimiento parte del conocimiento del valor de SINR (*Signal to Interference plus Noise Ratio*) medido en un canal. Además, se asume que existe una relación directa entre el SINR experimentado en la recepción de un bloque radio y el BLER (*Block Error Rate*) o probabilidad de error de bloque en dicha recepción para cada modo de transmisión.

Con todo esto, se podría calcular el *throughput* efectivo de la transmisión como:

$$Throughput = Rate(MCS) * \left(1 - BLER(MCS, SINR)\right) \tag{5}$$

donde Rate(MCS) es la tasa nominal del esquema de modulación y codificación MCS y BLER(MCS,SINR) es la tasa de error de bloque media (obtenida en base a simulación de nivel de enlace) para el esquema de modulación y codificación MCS y un valor de relación señal a interferencia más ruido media de bloque SINR.

La fórmula anterior es válida considerando un sistema sin retransmisiones o en el que las retransmisiones no se combinan. Además, en todo caso, se basa en la suposición de que la SINR medida será igual a la experimentada en la transmisión.

En cualquier caso se puede obtener una curva de máximo *throughput* en función de la SINR experimentada mediante la unión de segmentos de las curvas de *throughput* de los diferentes modos tomados en las regiones de SINR en las que dichos modos son óptimos. La figura Fig.4 muestra un ejemplo de curvas de *throughput* frente a SIR y de la curva de máximo *throughput*, que es la envolvente de las curvas individuales.

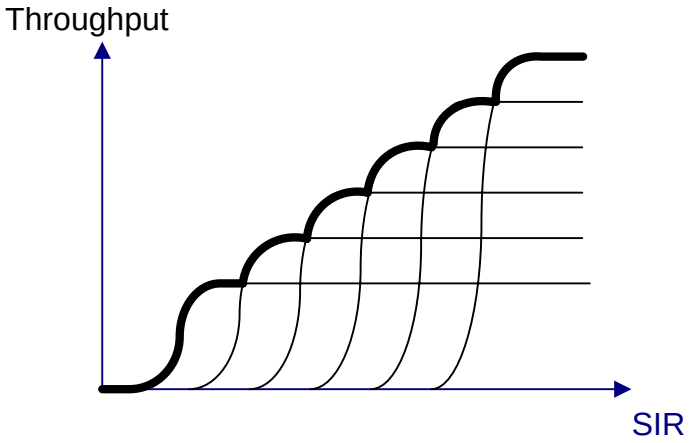

Fig.4. Curva de máximo *throughput* sin retransmisiones.

El esquema de *link adaptation* presentado muestra varias deficiencias. En primer lugar cabe destacar que se han asumido ciertas idealidades en torno a la SINR. La SINR medida inicialmente no es más que una estimación de la SINR experimentada real habiendo un error en la medida. Además, la SINR experimentada en la transmisión radio no es igual a la SINR medida inicialmente, debido a la naturaleza cambiante del canal radio y al retardo no despreciable que habrá siempre entre la medida inicial de la SINR y una transmisión radio adaptada a dicha SINR.

En segundo lugar la fórmula anterior no considera la combinación de retransmisiones típica de sistemas de comunicaciones móviles avanzados como EDGE o HSDPA. La capacidad de combinar retransmisiones permite mejorar mucho las prestaciones de ciertos modos, de forma que extienden la región de SINR en la cual proporcionan prestaciones aceptables e incluso la región de SINR en la que resultan óptimos.

El primer objetivo de este estudio genérico es proponer una estrategia de *link adaptation* que sí tenga en cuenta el error de estimación del SINR experimentado en la transmisión. Este error de estimación procede de dos fuentes: por un lado se supone la existencia de un error en la medida (acotado eso sí) y por otro lado existe un retardo no despreciable entre el instante en el que el móvil mide la SINR y el instante en el que recibe una transmisión adaptada al enlace en base al SINR medido y reportado. La estrategia propuesta en la subsección IV.2 no diferencia estos tipos de error sino que actúa teniendo en cuenta la estadística del error global.

Una vez realizado dicho estudio se propone una nueva fórmula de estimación del *throughput* efectivo que sí tiene en cuenta la variación de la SINR en las retransmisiones y la posible combinación. En base a dicha fórmula se extenderá la propuesta de la subsección IV.2 a una comunicación con retransmisiones.

IV.2 *Maximización del Throughput basada en pdfs del error de estimación de la SIR.*

En esta sección se presenta matemáticamente una estrategia para maximizar el *throughput* en un sistema con el LA genérico descrito en este apartado.

Se asume que se dispone de las funciones SINR-BLER para cada modo de transmisión. Además, se dispone de una función que relaciona el SINR medido y el SINR experimentado, de manera que para cada SINR medido, proporciona la densidad de probabilidad de percibir en recepción el SINR experimentado.

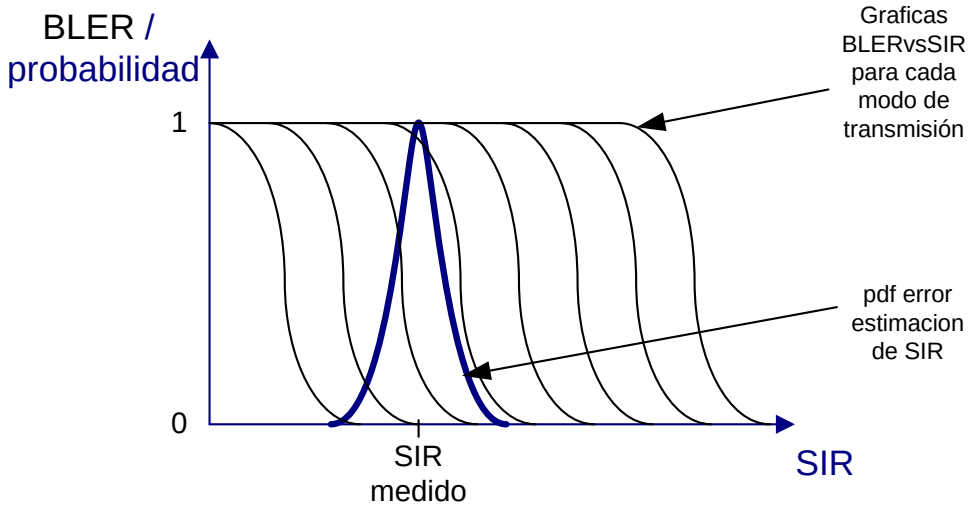

Fig.5 Maximización del *throughput*.

El *throughput* que experimentaría un usuario con un MCS dado, para un SINR medido concreto y conociendo la función densidad de probabilidad (fdp) del error de estimación de la SINR para ese valor de SINR medido, se obtendría con la siguiente formula:

$$Thr\left(MCS_n, SINR_{med}\right) = Rate\left(MCS_n\right)\left(1 - \int_{SINR_{\min}}^{SINR_{\max}} BLER\left(MCS_n, SINR\right)\ fdp\left(SINR_{med}, SINR\right) dSINR\right) \quad (6)$$

Dado un SINR medido, para maximizar el *throughput* de la transmisión bastaría con encontrar el MCS que maximice la formula anterior.

Con esta fórmula se consigue tener en cuenta el error en la estimación del SINR, incluyendo tanto el error en la medida como el producido por el retardo existente en el proceso de adaptación al enlace. Se puede obtener fácilmente una curva de máximo *throughput* frente a SINR medido, la cual expresa la capacidad de transmisión de un usuario en un sistema de forma más realista que la curva de máximo *throughput* sin errores en la estimación de SINR de la Fig.4.

IV.3 *Cálculo del throughput con retransmisiones.*

IV.3.1 *Throughput con una retransmisión.*

Se considera un protocolo en el que puede darse una primera transmisión y una retransmisión (segunda transmisión) para cada bloque. Si se definen t_1 y t_2 como la cantidad de primeras transmisiones y segundas transmisiones y $BLER_1$ como la tasa de error de bloque de las primeras transmisiones, se deberá cumplir que:

$$t_2 = BLER_1\ t_1 \tag{7}$$

A su vez, el número total de transmisiones t será la suma de t_1 y t_2, que se puede escribir como:

$$t = t_1 + t_2 = (1 + BLER_1)t_1 \tag{8}$$

Se pueden definir unos ratios de número de primeras transmisiones y segundas transmisiones con respecto a la cantidad total de transmisiones, $Ratio_{t_1}$ y $Ratio_{t_2}$ de la siguiente forma:

$$Ratio_1 = \frac{t_1}{t} = \frac{t_1}{(1 + BLER_1)t_1} = \frac{1}{1 + BLER_1} \tag{9}$$

$$Ratio_2 = \frac{t_2}{t} = \frac{BLER_1\ t_1}{(1 + BLER_1)t_1} = \frac{BLER_1}{1 + BLER_1} \tag{10}$$

Se puede obtener el *throughput* conseguido por las primeras transmisiones utilizando la misma fórmula empleada en el caso de no contar con retransmisiones (ec.(5)) pero multiplicando la tasa nominal del modo de transmisión por el ratio de primeras transmisiones y empleando el BLER de las primeras transmisiones:

$$Thr_1 = Ratio_1 Ratio_{1ok} Rate = \frac{1}{1 + BLER_1}(1 - BLER_1)\ Rate \tag{11}$$

De forma similar, el *throughput* conseguido en las retransmisiones es:

$$Thr_2 = Ratio_2 Ratio_{2ok} Rate = \left(\frac{BLER_1}{1 + BLER_1}\right)(1 - BLER_2)\ Rate \tag{12}$$

El *throughput* real será finalmente la suma de ambos:

$$Thr = Thr_1 + Thr_2 = \frac{1}{1 + BLER_1}(1 - BLER_1)\ Rate + \left(\frac{BLER_1}{1 + BLER_1}\right)(1 - BLER_2)\ Rate \tag{13}$$

IV.3.2 *Throughput con dos retransmisiones.*

De forma análoga al desarrollo anterior se puede obtener el *throughput* conseguido permitiendo dos retransmisiones.

Se definen $t_1, t_2,\ t_3$ y t como el número de primeras, segundas, terceras y totales transmisiones respectivamente. $BLER_1, BLER_2$ y $BLER_3$ son la tasa de error de bloque para las primeras, segundas (primeras retransmisión) y terceras (segunda retransmisión) transmisiones respectivamente. Con estas definiciones se tiene:

$$\begin{aligned} &t_1 \\ &t_2 = BLER_1\, t_1 \\ &t_3 = BLER_2 t_2 = BLER_2 BLER_1 t_1 \\ &t = t_1 + t_2 + t_3 = \left(1 + BLER_1 + BLER_1 BLER_2\right) t_1 \end{aligned} \tag{14}$$

De forma análoga al caso con una retransmisión se puede definir los ratios de primeras, segundas y terceras transmisiones con respecto al número total:

$$\begin{aligned} Ratio_1 &= \frac{1}{1 + BLER_1 + BLER_1 BLER_2} \\ Ratio_2 &= \frac{BLER_1}{1 + BLER_1 + BLER_1 BLER_2} \\ Ratio_3 &= \frac{BLER_1 BLER_2}{1 + BLER_1 + BLER_1 BLER_2} \end{aligned} \tag{15}$$

El *throughput* agregado será:

$$Thr = \left(Ratio_1 \underbrace{Ratio_{1-ok}}_{(1-BLER_1)} + Ratio_2 \underbrace{Ratio_{2-ok}}_{(1-BLER_2)} + Ratio_3 \underbrace{Ratio_{3-ok}}_{(1-BLER_3)} \right) Rate \tag{16}$$

IV.3.3 *Fórmula con N-1 retransmisiones.*

La fórmula del *throughput* agregado de la subsección anterior se puede extender fácilmente a un caso con N transmisiones (N-1 retransmisiones). Bastará con obtener una fórmula para los ratios de cada transmisión:

$$Thr = \left(\sum_{i=1}^{N} Ratio_i Ratio_{i-ok} \right) Rate \tag{17}$$

, donde:

$$\begin{aligned} Ratio_i &= \frac{\prod_{j=1}^{i} BLER_{j-1}}{\sum_{k=1}^{N} \left(\prod_{j=1}^{k} BLER_{j-1} \right)} \quad , con\ BLER_0 = 1 \\ Ratio_{i-ok} &= 1 - BLER_i \end{aligned} \tag{18}$$

Con la fórmula propuesta se pueden obtener nuevas curvas de *throughput* para cada modo con tal de que se disponga de la relación de BLER frente a SINR para la transmisión original y las sucesivas retransmisiones. Las nuevas curvas de *throughput* podrían tener una forma similar a la mostrada en la Fig.6.

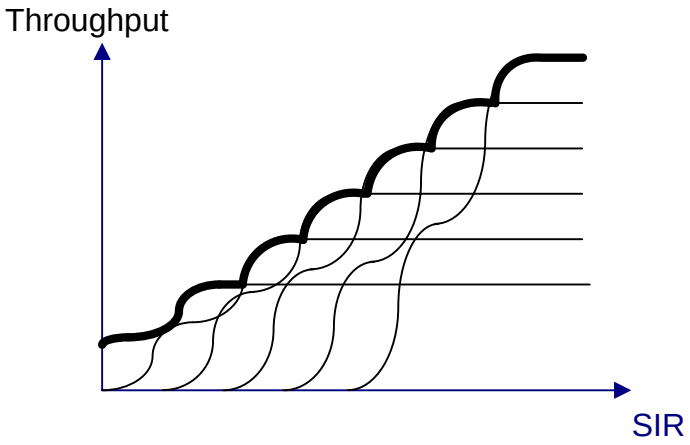

Fig.6 Curva de máximo *throughput* con una retransmisión.

IV.4 *Extensión del esquema de maximización del throughput con retransmisiones.*

Un análisis de la fórmula (6) lleva a la conclusión de que se trata de una variación de (5) sustituyendo el valor de BLER puntual para el $SINR_{med}$ por una expresión integral; este procedimiento se puede extender directamente para el caso con retransmisiones de modo que los valores de BLER desde la primera transmisión a la última se expresarían en forma integral. La única diferencia es que se precisarán tantas funciones de densidad de probabilidad del error de estimación de SINR como transmisiones se permitan.

Es decir, tendríamos una expresión como la mostrada en (17) donde:

$$BLER_0 = 1$$

$$BLER_i = \int_{SINR_{\min}}^{SINR_{\max}} BLER\left(MCS_n, SINR\right)\ fdp_i\left(SINR_{med}, SINR\right) dSINR \qquad (19)$$

, siendo $fdp_i(SINR_{med}, SINR)$ la función densidad de probabilidad del error de estimación de SINR para un $SINR_{med}$ concreto, para la i-ésima transmisión.

La justificación de porqué se utilizan diferentes fdp del error de estimación viene del diferente retardo existente entre la medida del SINR y la realización de la i-ésima transmisión. Este diferente retardo conlleva una diferente variación del SINR entre esos dos instantes y por lo tanto un diferente error de estimación.

V. ENTORNO DE SIMULACIÓN

En el presente estudio se ha utilizado un simulador dinámico a nivel de sistema del enlace descendente de HSDPA.

Se trata de un simulador por eventos discretos implementado en C++. En la actualidad forma parte de una plataforma de simulación de redes heterogéneas denominada SPHERE la cual ha dado lugar a varios artículos de investigación recientemente ([15][16]) en los que ha participado el autor

de esta tesina. Además, el simulador ha sido empleado en varios estudios sobre HSDPA (véanse y).

En este apartado se comentarán algunos de los modelos claves tenidos en cuenta en la simulación y que son importantes para la medida de prestaciones de un algoritmo de LA. A su vez, se presentarán los valores asumidos para los diferentes parámetros que configuran el escenario de simulación, justificando su elección.

Los modelos más relevantes para el presente estudio son los que hacen referencia a la caracterización del canal radio y su calidad.

El modelo de *path loss*, o de pérdidas de propagación, empleado se ha extraído de [19]:

$$L_p(dB) = \left(40 - 16 \cdot 10^{-2} \Delta h_b\right) \log_{10} d - 18 \log_{10} \Delta h_b + 21 \log_{10} f + 80 \qquad (20)$$

,donde d es la distancia en km entre la estación base y el móvil, f es la frecuencia portadora en MHz y Δh_b es la altura de la antena de la estación base sobre el nivel de los edificios medida en metros. Empleando los valores $\Delta h_b = 15$ y $f = 2000$ se obtiene la fórmula recogida en la Tabla 4.

El *shadowing*, o desvanecimiento a largo plazo, se ha modelado e implementado como una variable lognormal con correlación espacial. Se ha seguido el modelo de [20] y la implementación de [21]. La distancia de correlación considerada ha sido de 20m y la desviación estándar del *shadowing* 8dB, que son valores comunes en la bibliografía.

El *fast fading*, o desvanecimiento a corto plazo, se ha modelado de acuerdo a lo propuesto en [22], documento realizado en el marco de la Red de Excelencia Europea Newcom (*Network of Excellence in Wíreless Communications)* en la cual participa el grupo investigador del autor de esta tesina. En este modelo se define matemáticamente la función densidad de probabilidad de un factor de *fast fading*. Dicho factor multiplica a la potencia recibida en el enlace. Particularizando la fórmula propuesta en [22] para el caso de 3 caminos independientes resolubles en el receptor, se tiene la fórmula recogida en la Tabla 4.

El factor de ortogonalidad (véase [23] para una explicación del mismo) se ha considerado fijo en toda la simulación, pero se ha diferenciado según la velocidad del usuario. Los valores escogidos han sido de 0.9 para usuarios peatonales (3 km/h) y de 0.5 para usuarios vehiculares (50 km/h). Estos valores están en consonancia con la fórmula aproximada del factor de ortogonalidad comentada en [23] y los modelos de canal pedestrian A y vehicular A de [19].

El modelo de interferencias es idéntico al empleado en [17]. En dicho modelo se asume que el receptor no es capaz de cancelar la interferencia producida entre los diferentes códigos de su señal útil. Este modelo es común en las investigaciones realizadas por grupos punteros en la materia[1].

Las relaciones entre SINR-BLER, que suponen la interfaz entre la simulación a nivel de sistema y a nivel de enlace, han sido obtenidas de [22]. Estas relaciones asumen que en caso de haber retransmisiones se produce combinación de las mismas según el esquema de Chase Combining lo cual ha impone la restricción de tener que utilizar este tipo de combinación en el estudio.

El modelo de tráfico más empleado en los estudios llevados a cabo y que se referirá como de 'buffer lleno' asume que la estación base siempre tiene datos para transmitir al terminal móvil, y que dichos datos siempre son suficientes como para llenar el bloque radio más grande de datos que se pueda transmitir.

Para este estudio vamos a suponer la existencia de 30 modos diferentes de transmisión en HSDPA, concretamente tomaremos los 30 modos asociados a los CQIs para la categoría 10 de móviles HSDPA.

La Tabla 4 recoge los principales modelos empleados en el simulador junto con los valores dados a los parámetros más importantes.

Parámetro	Modelo / Valor
Potencia transmitida por el Nodo-B	43 dBm
Potencia del canal piloto (CPICH)	33 dBm
Pathloss	$L(dB) = 128.1 + 37.6 \log_{10} d(Km)$
Shadowing	Distribución lognormal, 8dB desviación estándar, distancia de correlación 20m.
Fast fading	$f_{\psi}(\psi) = \frac{27}{2}\psi^2 e^{-3\psi}$
Velocidad	Usuarios peatonales : 3 Km/h, usuarios vehiculares: 50 Km/h
Factor de ortogonalidad	Usuarios peatonales: 0.9, usuarios vehiculares: 0.5
Radio celular	500 m
Número de celdas	7
Número de celdas interferentes	6
Error medida CQI (SINR)	Distribución lognormal, media 0, desviación estándar 1dB
Categoría de Terminal	10
Número de procesos HARQ	6
Número de transmisiones permitidas	variable
Parámetros del bucle externo para LA	$A_{up} = 9*A_{down,}$ para BLER objetivo = 10%
Modos de transmisión	30 modos asociados a los CQI de la categoría 10
Tipo de combinación	Chase combining

Tabla 4: Parámetros de simulación

VI. ESTUDIO DEL LINK ADAPTATION EN HSDPA.

VI.1 *Análisis del algoritmo común sin retransmisiones.*

A partir de las curvas de *throughput* frente a SINR de los modos HSDPA empleados se podrían obtener valores de *throughput* o SINR umbrales que delimitarían las regiones de trabajo de los diferentes modos, como se puede observar en la Fig.7.

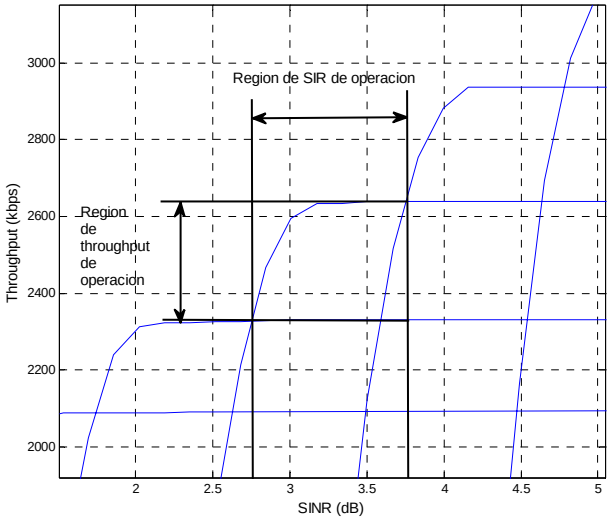

Fig.7. Curvas de *throughput* y umbrales entre modos.

Para cada modo se podría definir, además, una región de BLER de trabajo, estableciendo unos BLER umbrales máximo y mínimo, tal y como ya se expuso para el caso de GPRS en la subsección III.1. Los valores de BLER umbrales para el modo n-ésimo, pueden calcularse fácilmente con las fórmulas dadas en la sección III.1, donde $Thr_{n,\max}$ se sustituye por la tasa nominal del modo n-ésimo ($Rate_n$), mientras que $Thr_{n,\inf}$ y $Thr_{n,\sup}$ se sustituyen por las tasas del modo n-ésimo y (n-1)-ésimo (dado que el *throughput* umbral inferior coincide con la tasa del modo inferior y el *throughput* umbral superior coincide con la tasa del modo superior). De este modo, los umbrales máximo ($BLER_{n,\max}$) y mínimo ($BLER_{n,\min}$) quedan:

$$BLER_{n,\max} = 1 - \frac{Thr_{n,\inf}}{Thr_{n,\max}} = 1 - \frac{Rate_{n-1}}{Rate_n} \tag{21}$$

$$BLER_{n,\min} = 1 - \frac{Thr_{n,\sup}}{Thr_{n,\max}} = 1 - \frac{Rate_n}{Rate_n} = 0 \tag{22}$$

El *link adaptation* óptimo (en términos de *throughput*) debería elegir el modo de transmisión más alto para el cual se garantizara un BLER dentro del margen de BLER del modo. En este caso basta con el que el BLER sea inferior al máximo, puesto que en HSDPA el mínimo es 0. Dado que la relación de tasa entre modos asociados a CQIs consecutivos no es fija y que dicha relación

determina los BLER umbrales tal y como se recoge en la ecuación (21) y (22), queda claro que no existe un BLER umbral genérico para realizar el *link adaptation*. Si bien la definición de un BLER genérico sí puede ser impuesta por cuestiones de calidad de servicio (para evitar mayores retardos de los necesarios), no daría lugar a un algoritmo óptimo en cuanto a la maximización del *throughput*.

Como ya se ha visto, en HSDPA el terminal móvil es el encargado de informar de forma periódica a la estación base de la calidad radio que percibe. Concretamente, el móvil le indica al Nodo-B en forma de CQI (*Channel Quality Indicator*) el modo de transmisión que podría recibir con un BLER menor al 10%. Por lo tanto, al fijarse un BLER objetivo genérico no óptimo, un mecanismo de *link adaptation* en el que el *scheduler* del Nodo-B eligiera para la transmisión el modo indicado por el CQI sería sub-óptimo en términos de *throughput*. No obstante, al ser la caída de la función BLER para cada modo muy abrupta (véase la Fig. 3), la diferencia en los umbrales óptimos y los obtenidos para BLER del 10% no es muy significativa y en consecuencia la variación en *throughput* tampoco lo debería ser.

Hay otros problemas más perniciosos para las prestaciones del sistema como el error en la medida del CQI. Generalmente se asume que dicho error presenta una distribución lognormal de media nula y 1 dB de desviación estándar, con lo que no merece la pena hilar muy fino en la elección del BLER umbral considerando diferencias menores a 1 dB si luego la medida del CQI no es tan precisa. Por tanto, pese a que los medios puestos en el estándar para el *link adaptation* no son óptimos, en principio, para maximizar el *throughput*, dadas las circunstancias son aceptables.

Además del error en la medida, existe un importante error debido al retardo en el *link adaptation*. Canales muy variables combinados con retardos grandes pueden llevar a una estimación de calidad del canal basada en el CQI muy imprecisa y, por tanto, pueden producir muchas elecciones del modo de transmisión no óptimas: elecciones conservativas u optimistas.

La solución propuesta en la bibliografía para el problema de los errores cometidos en la estimación del canal y presentada en la subsección III.4.1 es similar a la adoptada en el control de potencia de bucle externo que empleaba UMTS convencional. Se trata de corregir la información proveniente de los CQIs tratando de garantizar un cierto BLER en la transmisión.

En la Fig.8 se puede observar la CDF (función densidad de probabilidad acumulada) del *throughput* de un usuario peatonal a 3 km/h, calculado en intervalos de 1 segundo. Se presentan, por un lado, dos casos ideales en los que se supone que el Nodo-B conoce exactamente la SINR experimentada por el terminal (o la SINR más ruido) y transmite la información según los umbrales estándar o según los umbrales optimizados. En este caso, el simulador se ha empleado para obtener series de SINR medidas por un usuario moviéndose por la celda sin errores (‘ideal’) o con errores de media (‘solo error de medida’). A continuación, dicho SINR se ha traducido a un valor de CQI empleando unos valores de SINR umbrales y dicho CQI en un valor de *throughput* efectivo empleando el SINR real y las tablas de SINR-BLER. Por último, los valores de *throughput* efectivo

han sido agregados en intervalos de 1 segundo y finalmente se ha calculado la CDF de esta magnitud agregada. Además, se incluyen las curvas para el caso sin errores en la medida pero retardo realista en el *link adaptation* (‘solo error de retardo’), para el caso de retardo realista y error en la medida sin corrección (‘realista’) y, por último, para el caso con retardo realista, error de medida y corrección (‘realista + bucle de corrección’). Los resultados presentados por estas tres últimas curvas se han obtenido exclusivamente con el simulador de sistema. Se puede apreciar claramente que en el caso de un usuario peatonal la degradación de prestaciones viene dada sobretodo por el error en la medida, que entre otros efectos limita el máximo *throughput* alcanzable, más que por el error debido al retardo en el proceso de *link adaptation*. Se aprecia una considerable mejora de prestaciones al añadir el bucle de control; este mecanismo consigue unas prestaciones más cercanas al caso sin error en la medida excepto que no se alcanzan los *throughputs* más altos.

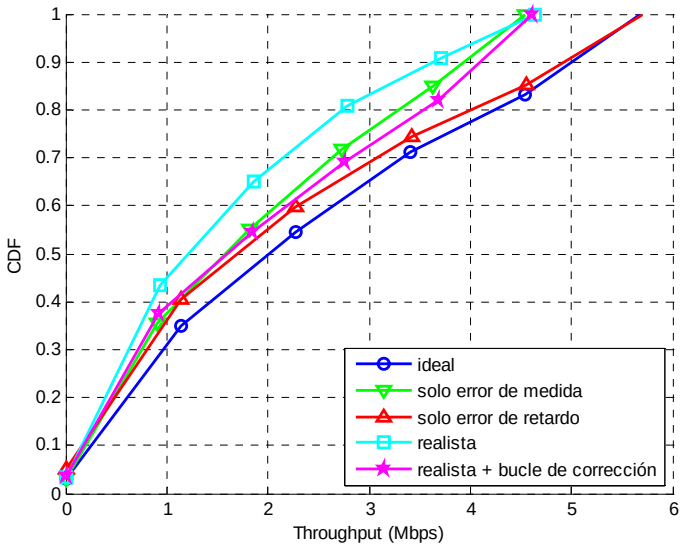

Fig.8. CDF del *throughput* para un usuario peatonal sin retransmisiones.

Las diferencias entre los escenarios son más visibles si comparamos los valores de *throughput* medio en kbps:

Ideal	Solo error de medida	Solo error de retardo	Con error realista	Realista y bucle externo
2336.0	1813.2	2128.2	1523.7	1877.2

Tabla 5: Throughput medio en kbps para un usuario peatonal sin retransmisiones.

La mejora al usar el bucle de corrección con respecto al caso con error es del orden del 23%. Aún se podría esperar cierta mejora si se consiguieran alcanzar *throughputs* parciales altos.

En la Fig.9 se muestra la CDF del *thoughput* de un usuario con velocidad vehicular a 50 Km/h, para cinco escenarios idénticos a los estudiados para el usuario peatonal. Una comparación de las

curvas lleva a la conclusión de que el error por el retardo del proceso degrada el *throughput* aumentando mucho la probabilidad de que se den tasas bajas (comparar curva 'ideal' con la de 'solo error de retardo'), mientras que la existencia del error en la medida produce una clara limitación en el máximo *throughput* alcanzable (comparar curva 'ideal' con la de 'solo error de medida'). La curva realista muestra ambos efectos perniciosos. Se aprecia una leve mejora de prestaciones al añadir el bucle de control.

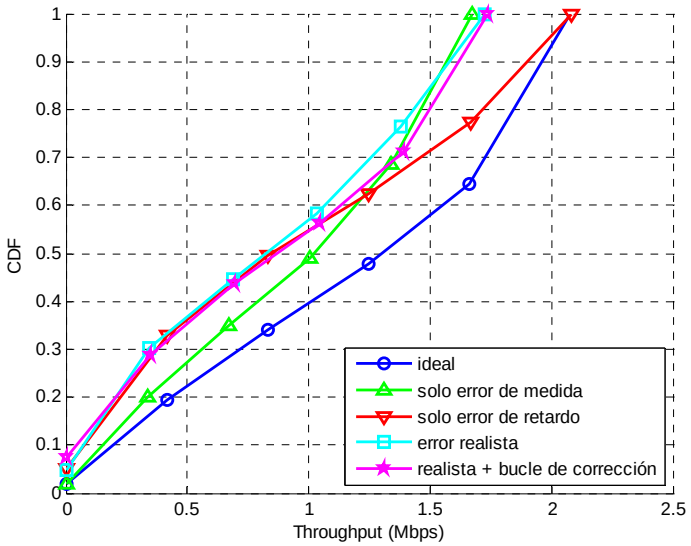

Fig.9. CDF del *throughput* para usuario vehicular sin retransmisiones.

Las diferencias entre los escenarios son más visibles si comparamos los valores de *throughput* medio en Kbps:

Ideal	Solo error de medida	Solo error de retardo	Con error realista	Realista y bucle externo
1203.7	998.9	954.6	817.2	864.6

Tabla 6: Throughput medio en kbps para un usuario vehicular sin retransmisiones.

La mejora al usar el bucle de corrección con respecto al caso con error es del orden del 6%. Si bien, es una mejora significativa, al comparar el valor obtenido con los casos ideales se intuye que una mayor mejora es posible. Es de notar, que persiste la limitación en el máximo *throughput* alcanzable como en el caso de usuarios peatonales.

La intuición expresada y que también se mostró en el caso de usuarios peatonales viene reforzada por varias cuestiones. Por un lado ya se ha visto que tratar de garantizar un cierto BLER es subóptimo en términos de *throughput* (siendo una solución de compromiso en ese caso tener un BLER de 0.1). Por otro lado, el método de corrección se realiza teniendo en cuenta la información proveniente de los reconocimientos de datos enviados por el receptor, pero todo el procesado

necesario de estos reconocimientos implica un retardo en la adaptación al enlace que pudiera ser intolerable. Además, el modo en el que se fuerza a tener un BLER dado es artificial e ineficiente. Piénsese por ejemplo en un enlace muy estable que sin bucle cerrado podría fijar su modo de transmisión y no presentar errores, mientras que con el bucle cerrado presentará errores al forzarse el BLER de 0.1. Por todo ello se cuestiona que este algoritmo sea óptimo y se propondrán alternativas más adelante.

Tanto en el estudio anterior como en los posteriores se asume un valor de A_{down}=0.05 y A_{up}=0.45 en el algoritmo de corrección del CQI. Esta configuración ha mostrado ser óptima al ser comparada con otras posibles.

VI.2 *Análisis del algoritmo común con retransmisiones.*

En todos los escenarios estudiados hasta el momento se ha supuesto que no hay retransmisiones. La existencia de retransmisiones combinables puede, en ocasiones, mejorar considerablemente las prestaciones de un sistema en cuanto al *throughput* ofrecido.

La Fig.10 presenta las CDF del *throughput* de un usuario peatonal con y sin retransmisiones para los escenarios con error realista sin corrección (‘error realista’) y con corrección (‘error realista + bucle’). La realización de retransmisiones mejora mucho las prestaciones del sistema minimizando el efecto del error en el *link adaptation*. En este caso, la adición del bucle de control reduce el porcentaje de retransmisiones del 18% al 10%, pero es ineficiente y resulta perjudicial para el *throughput*, como se puede observar en la Fig.10 y la Tabla 7.

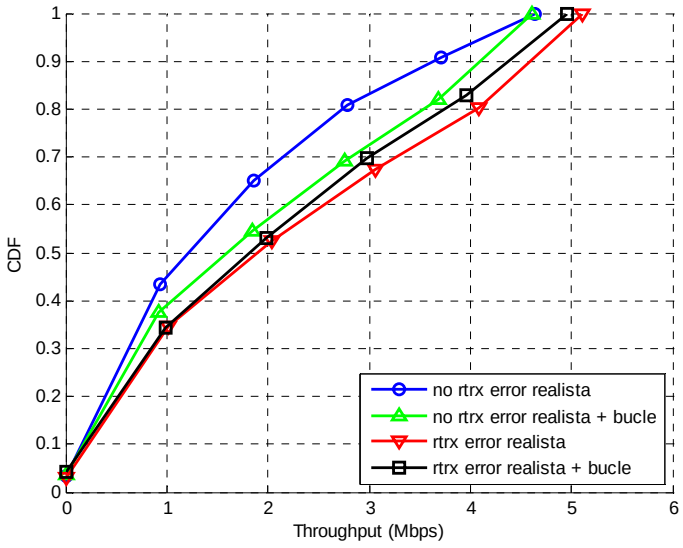

Fig.10. CDF del *throughput* para usuario peatonal.

Sin rtrx, con error realista	Sin rtrx, con error realista y bucle	Con rtrx, con error realista	Con rtrx, con error realista y bucle
1523.7	1877.2	2177.8	2060.6

Tabla 7: Throughput medio en kbps para un usuario peatonal.

La Fig.11 presenta los resultados para los mismos escenarios que el caso anterior, pero para un usuario vehicular. De nuevo, la realización de retransmisiones mejora las prestaciones del sistema aunque no tanto como ocurría en el caso peatonal (las gráficas están mucho más cercanas). La adición del bucle de corrección reduce el porcentaje de retransmisiones del 25% (hay más errores originalmente que en el caso anterior por la mayor variación del canal del usuario vehicular) al 10%, pero vuelve a ser ineficiente y reduce el *throughput*, como se puede observar en la Fig.11 y Tabla 8.

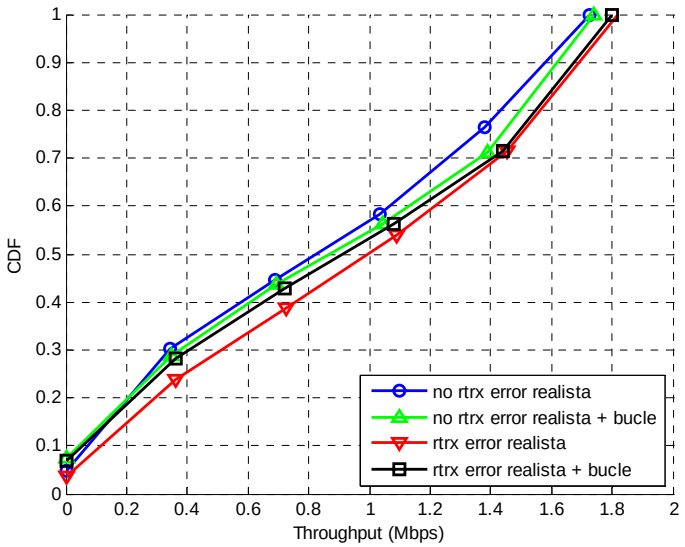

Fig.11. CDF del *throughput* para usuario vehicular.

Sin rtrx, con error realista	Sin rtrx, con error realista y bucle	Con rtrx, con error realista	Con rtrx, con error realista y bucle
817.2	864.6	952.1	899.5

Tabla 8: Throughput medio en kbps para un usuario vehicular.

En conclusión, la adición del bucle de control si bien es positiva en ausencia de retransmisiones, en presencia de éstas resta eficiencia a la comunicación para los escenarios estudiados. Dado que las velocidades estudiadas son comunes en un escenario urbano, y que éste es el mercado fundamental de HSDPA se desaconseja el uso del bucle de control.

VI.3 *Otras técnicas de procesado de los CQIs.*

Partiendo de un análisis profundo de los mecanismos que producen el error de estimación en el CQI, se propuso prefiltrar o procesar los informes de CQI llegados al Nodo-B antes de ejecutar la adaptación al enlace, trabajo publicado en [17].

En el citado artículo se asumió una medida de SINR libre de errores y una comunicación sin retransmisiones. A continuación se realiza una evaluación más realista, suponiendo errores de medida y la posibilidad de realizar retransmisiones. Como técnica de procesado se ha elegido la predicción con el método LMS (*Least Mean Squares*) de la serie de CQIs filtrados con un filtro promediador de longitud 64 muestras. Se ha elegido este procesado por ser el óptimo en [17]. Además se propone una modificación a lo estudiado en [17] consistente en aplicar el procesado de los CQIs y posteriormente el bucle de corrección.

En la Fig.12 se observa que, en el caso de un usuario peatonal, el procesado no aporta mejora por sí mismo, pero combinado con el bucle puede mejorar las prestaciones de éste último aislado. La mejora se aprecia en el incremento del valor de *throughput* máximo alcanzable.

En la Fig. 13 se observa que la aplicación del procesado propuesto es muy beneficiosa para un usuario vehicular, mucho más que el empleo simple del bucle de corrección. La aplicación conjunta de ambos proporciona unas prestaciones ligeramente inferiores a la del procesado LMS para *throughputs* altos.

Por tanto, la aplicación conjunta del filtrado y el bucle de control ha demostrado ser una solución válida tanto para usuarios vehiculares como peatonales, mejorando siempre la prestaciones del bucle de control trabajando aislado. El hecho de que la mejora sea válida para ambas velocidades evita que el sistema tenga que distinguir a los usuarios por velocidad, que era uno de los problemas que planteaba en [17] donde se empleaba únicamente el procesado.

Con error realista	Con error realista y bucle	Con error realista y procesado	Con error realista, bucle y procesado
1523.7	1877.2	1847.2	2209.8

Tabla 9: Throughput medio en kbps para un usuario peatonal sin retransmisiones.

Con error realista	Con error realista y bucle	Con error realista y procesado	Con error realista, bucle y procesado
817.2	864.6	1038.0	988.7

Tabla 10: Throughput medio en kbps para un usuario peatonal sin retransmisiones.

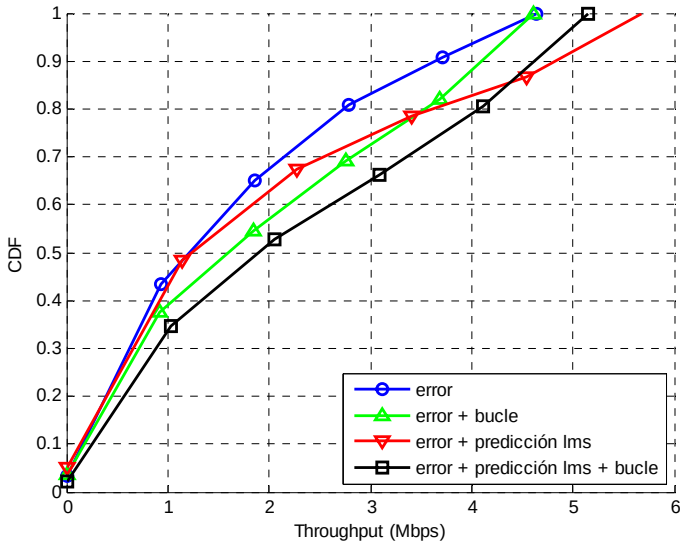

Fig.12 CDF del *throughput* para usuario peatonal sin retransmisiones.

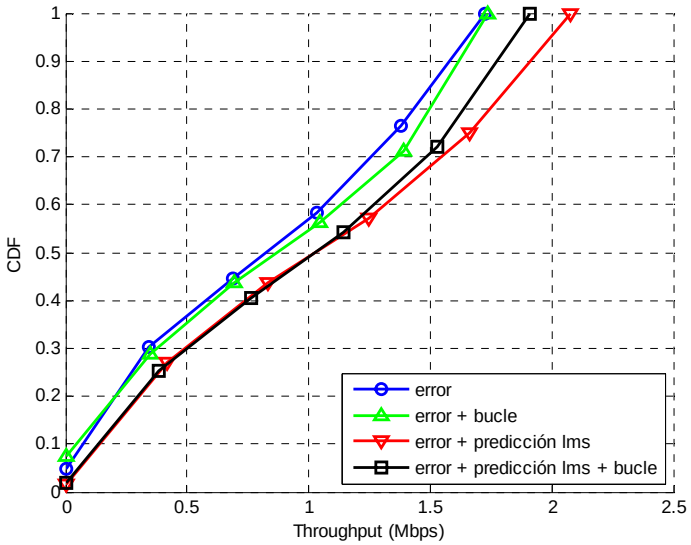

Fig.13 CDF del *throughput* para usuario vehicular sin retransmisiones.

Al realizar retransmisiones, de nuevo la estrategia combinada se muestra como una opción válida. En la Fig. 14, para un usuario peatonal se observa que la predicción LMS por sí misma no añade ninguna mejora, pero en su operación conjunta con el bucle es capaz de mejorar el funcionamiento de este último operando aislado y acercarse a las prestaciones de un caso sin procesado. Aún así se produce cierto empeoramiento de prestaciones con respecto a un sistema sin corrección.

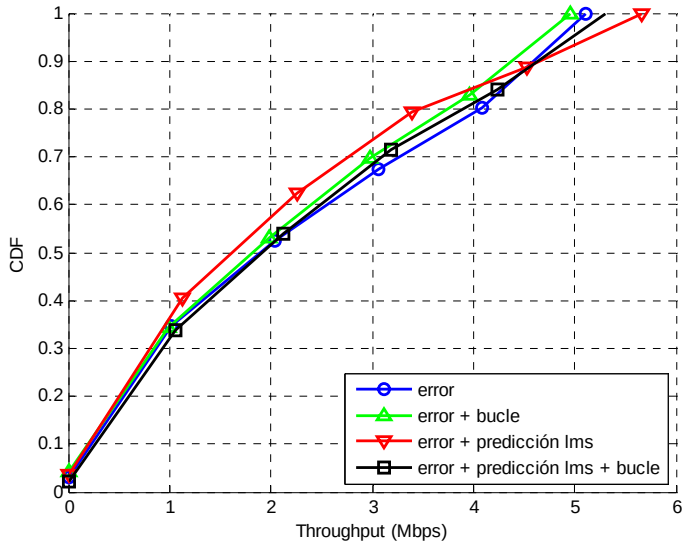

Fig.14 CDF del *throughput* para usuario peatonal con retransmisiones.

Con error realista	Con error realista y bucle	Con error realista y procesado	Con error realista, bucle y procesado
2177.8	2060.6	1984.7	2173.0

Tabla 11: Throughput medio en kbps para un usuario peatonal con retransmisiones.

La situación es diferente para un usuario vehicular. En la Fig. 15, se observa que en ese caso tanto la predicción LMS como la operación conjunta de predicción LMS y bucle de corrección introducen una mejora en las prestaciones del sistema.

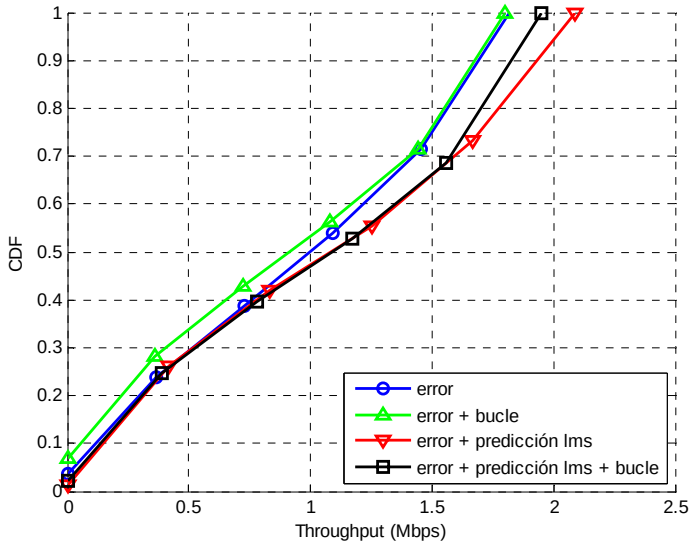

Fig.15 CDF del *throughput* para usuario vehicular con retransmisiones.

Con error realista	Con error realista y bucle	Con error realista y procesado	Con error realista, bucle y procesado
952.1	899.5	1071.0	1040.0

Tabla 12: Throughput medio en kbps para un usuario vehicular con retransmisiones.

A la luz de los resultados se reafirma la utilidad del método de operación conjunta de procesado y bucle de corrección. Ya que en el caso peatonal es mejor que usar únicamente el bucle, mientras que introduce sólo una ligera degradación de prestaciones con respecto al caso sin corrección, y en el caso vehicular es claramente positivo.

VI.4 *Maximización del Throughput mediante distribución del error de estimación del CQI.*

En esta subsección se aplicará al caso concreto de HSDPA la teoría desarrollada para el *link adaptation* genérico. Se hará especial hincapié en que las estrategias adoptadas sean realizables y muestren una mejora de prestaciones en un amplio conjunto de escenarios de operación de HSDPA.

VI.4.1 *Planteamiento del método.*

Obtener funciones que para cada SINR medido proporcionen la fdp del error de estimación de la SINR experimentada es prácticamente irrealizable en un escenario real.

Una estrategia posible pasaría por asumir una cuantificación de los valores de SINR. Para cada SINR cuantificado medido se podría disponer de una distribución de probabilidad de la ocurrencia del SINR cuantificado experimentado.

En HSDPA, el Nodo-B no dispone de informes frecuentes del SINR pero sí de una magnitud relacionada, el CQI. Esta magnitud está cuantificada pudiendo tomar 31 valores diferentes facilitándose así la obtención de las distribuciones de probabilidad de ocurrencia mencionadas. Se puede realizar la aproximación consistente en que, dado un informe de CQI medido, el informe recibido 3 subtramas después (6ms) equivale al CQI experimentado asociado al CQI medido.

Las distribuciones de probabilidad de ocurrencia se podrían compactar en una estructura de 31x31 valores, con 31 distribuciones (tantas como posibles CQI medidos) de 31 puntos (31 posibles CQI experimentados). La Fig.16 presenta una hipotética distribución del CQI experimentado para un CQI medido de valor 10.

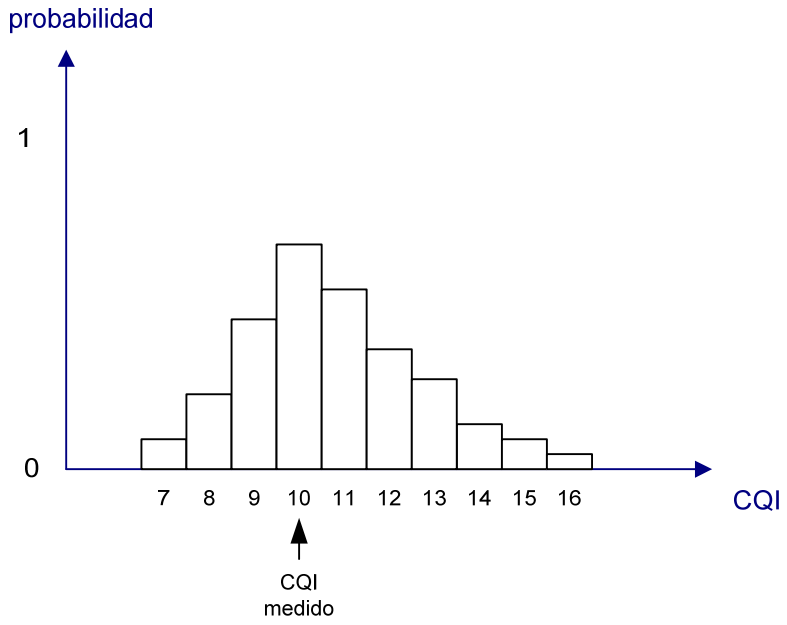

Fig.16 Ejemplo de distribución de ocurrencia de un CQI experimentado dado un CQI medido.

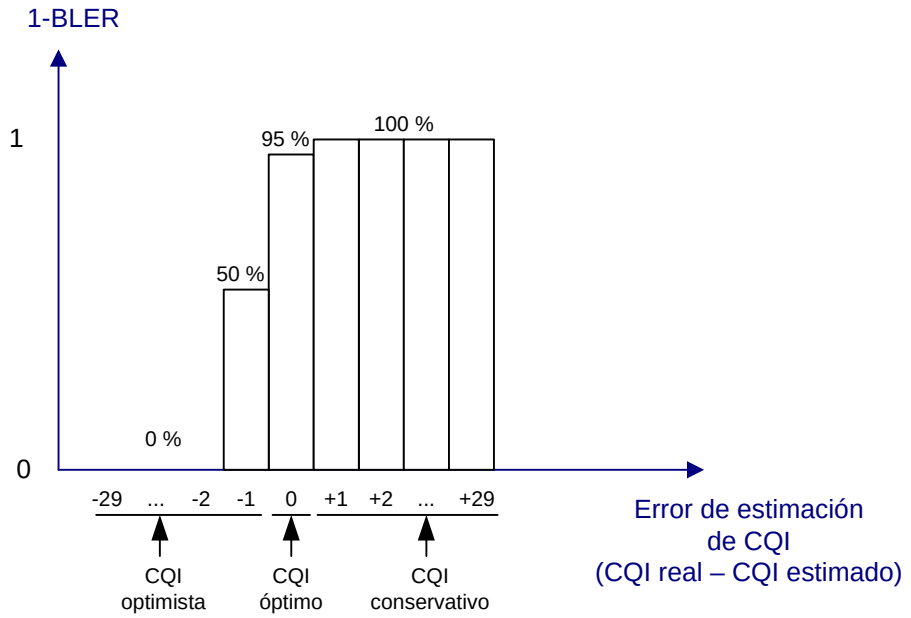

Fig.17 Función 1-BLER frente a error de estimación de CQI (CQI_{real}-$CQI_{estimado}$).

Junto con estas distribuciones sería necesario disponer de una función que relacionase el error de estimación del CQI con un valor de BLER. La siguiente gráfica recoge esta información. En los casos en los que el CQI medido coincide con el CQI experimentado el error de estimación es 0, y se asume un BLER del 5% intermedio entre el 10% y el 0% que se tienen en la frontera que delimita la asignación de SINR a CQI. En caso de que el CQI experimentado sea mayor que el medido, el error de estimación es negativo y viendo las gráficas de SINR-BLER se deduce un valor de 1-BLER del 100%. Por otro lado, en caso de CQI menor al experimentado se distinguen dos casos. Si la diferencia es de 1, la transmisión experimenta un BLER de entre el 10% y el 90%, tomando un valor intermedio de BLER tenemos un 1-BLER del 50%. En cambio, si la diferencia es

mayor de 1, se asume que la transmisión se recibe siempre mal y 1-BLER es del 0%. La Fig.17 recoge esta reflexión.

Dado un CQI medido, el algoritmo de *link adaptation* ejecutado en el Nodo-B puede realizar la transmisión de datos según el modo asociado a dicho CQI o a otro CQI cualquiera. Para maximizar el *throughput* habría que maximizar el producto de la tasa del modo elegido por la proporción de bloques del modo elegido correctamente recibidos en recepción en las condiciones determinadas por el CQI medido. Esta proporción consiste en la suma del producto de la distribución de ocurrencia para el CQI medido por la función 1-BLER centrada en el CQI elegido.

Por ejemplo, si el CQI medido es 10 y el CQI propuesto para ser elegido (CQI estimado) es 11 habrá que multiplicar la distribución de error para el CQI medido 10 (ya presentada anteriormente) por la función 1-BLER centrada en el 11, tal y como se muestra en la Fig.18.

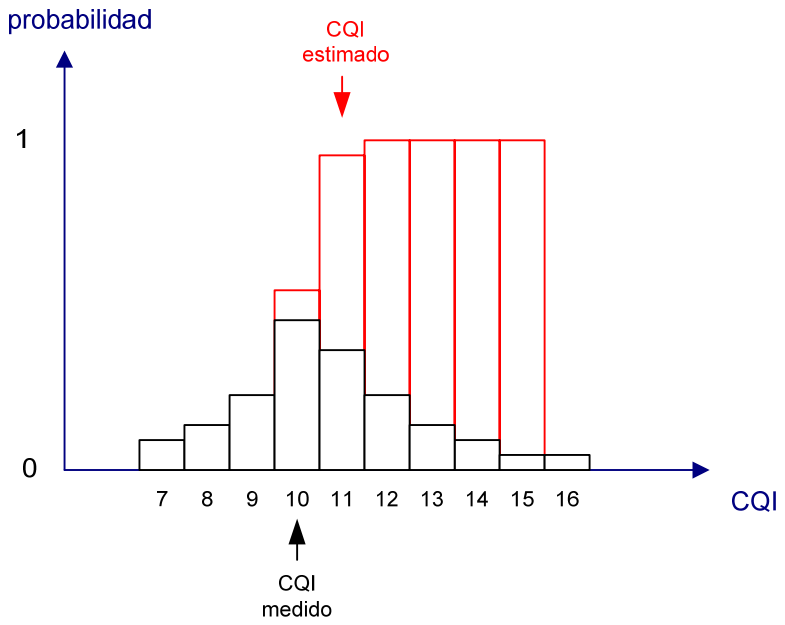

Fig.18 Representación del producto de funciones empleado en el método.

La suma del producto de estas dos funciones se multiplica por la tasa del modo 11 resultando en el *throughput* medio esperado. La siguiente fórmula recoge esta explicación:

$$Thr(CQI_{med}, CQI_{est}) = Rate(CQI_{est}) \left(\sum_{CQI=1}^{30} f_{1-BLER}(CQI - CQI_{est})\ pdf\,(CQI_{med}, CQI) \right) \tag{23}$$

,donde CQI_{med} y CQI_{est} son el CQI medido y el estimado respectivamente, $Thr(CQI_{med}, CQI_{est})$ es el *throughput* medio esperado para el CQI medido si se eligiera el CQI estimado, $Rate(CQI_{est})$ es la tasa nominal del CQI estimado, $f_{1-BLER}(CQI - CQI_{est})$ es la función que relaciona el error de estimación con la probabilidad de recibir correctamente un bloque del modo CQI estimado y $pdf\,(CQI_{med}, CQI)$ es la distribución de probabilidad de ocurrencia de los diferentes CQI dado un valor del CQI medido.

El CQI óptimo, el elegido para transmitir, es aquel CQI estimado que maximiza el *throughput* medio y, por tanto, la expresión (23):

$$CQI_{elegido} = \arg\max_{CQI_{est}} \; Thr(CQI_{med}, CQI_{est}) \tag{24}$$

VI.4.2 *Comentarios y modificaciones propuestas.*

El mecanismo propuesto tiene aún un problema a solventar. En HSDPA todos los CQIs recibidos en la estación base están corrompidos con ruido de medida con lo que no está asegurado que la maximización anterior sea óptima. Aún así, la estadística de este ruido permite emplear métodos que lo eliminen parcialmente. Se evaluará el método con y sin eliminación de ruido.

El método, por otro lado, sólo será óptimo si el retardo en el link adaptation fuera fijo. Ese caso de retardo fijo sólo se cumple si el informe del CQI es periódico y con periodo 2ms, en cuyo caso el retardo es de unos 7ms. Pero si el periodo es mayor, por ejemplo x ms, el retardo en la estimación del CQI puede ser desde 7 hasta 7+(x-2)ms donde x es múltiplo de 2ms. Habrá que tener distribuciones de ocurrencia del CQI para cada retardo posible, lo cual para periodos de CQI grandes puede ser inabordable pero no para periodos pequeños. Una opción es modelar el efecto de ese aumento de retardo. Puede que se mantenga la forma general de la distribución pero sufra un ensanchamiento. En este estudio no se ha tenido en cuenta esta cuestión ya que lo más común es que las redes se configuren de modo que el CQI se envíe cada 2ms.

Además, se suscita la duda de si habría que tener distribuciones de ocurrencia del CQI diferentes según la movilidad de los usuarios, la carga del sistema, etc. Esta cuestión se dilucidará en la siguiente subsección en base a simulaciones del funcionamiento del método.

Por último, una cuestión importante es la posibilidad de que haya retransmisiones. En ese caso, la expresión del *troughput* seguiría la expresión de (17), y de manera análoga a (19) se emplearían los valores de BLER obtenidos de la siguiente forma:

$$BLER_0 = 1$$

$$BLER_i = 1 - \left(\sum_{CQI=1}^{30} f_{1-BLER}(CQI - CQI_{est}) \; pdf_i(CQI_{med}, CQI) \right) \tag{25}$$

siendo $pdf_i(CQI_{med}, CQI)$ la distribución de probabilidad de ocurrencia de un CQI dado conociéndose el CQI medido, para la i-ésima transmisión. La obtención de la distribución para la primera transmisión es trivial como ya se expuso. Pero para obtener la distribución del resto de transmisiones es necesario tener en cuenta la combinación de las retransmisiones. A continuación se explica brevemente cómo habría que realizar el cálculo. Se asume que un terminal informa de un CQI_{med}, 6ms más tarde informa de un CQI_1 y 12ms después informa de un CQI_2. La distribución para la primera transmisión tiene en cuenta la probabilidad de ocurrencia de los diferentes CQI_1 dado un CQI_{med}, pero la distribución para la segunda transmisión debería recoger la probabilidad de ocurrencia de un CQI efectivo combinación de CQI_1 y CQI_2. De forma semejante ocurriría si se

tuvieran en cuenta más retransmisiones. Para calcular el CQI efectivo podrían traducirse los valores de CQI parciales en valores de SINR. Los SINR resultantes se sumarían en unidades lineales y se volverían a traducir a un CQI.

La estrategia basada en retransmisiones no ha dado lugar hasta el momento a una mejora en *throughput* de ahí que se omitan sus resultados. Se está estudiando el preprocesado de los CQIs previamente a la obtención de las distribuciones para eliminar ruido en las series de CQIs y mejorar las prestaciones de la técnica. Esta línea de trabajo ya se ha mostrado útil en el caso sin retransmisiones como se muestra en la siguiente subsección.

VI.4.3 *Pruebas del método.*

En esta sección se aportan los resultados de aplicar el método propuesto en simulaciones a nivel de sistema. Previamente se han simulado usuarios en movimiento y se han obtenido las distribuciones de ocurrencia del CQI. Concretamente se dispone de distribuciones particularizadas para usuarios peatonales, distribuciones particularizadas para usuarios vehiculares y distribuciones generalizadas resultantes del promedio de las anteriores.

Empleando dichas distribuciones se simula el movimiento de un usuario peatonal y otro vehicular y se comprueban las prestaciones de cada método comparándolos con algunos de los estudiados en la sección anterior.

Como se puede comprobar en la Tabla 13, para el usuario peatonal, el hecho de emplear el método basado en distribuciones del error proporciona resultados intermedios entre el peor caso (filtrado simple) y el mejor (bucle de corrección más filtrado). Sorprende las buenas prestaciones para la distribución global y queda claro que la distribución vehicular no es apta.

Para el usuario vehicular, de nuevo se obtiene una mejora sobre el caso peor (sin corrección), sin llegar a las prestaciones de los algoritmos óptimos (filtrado simple y filtrado más bucle de corrección). Sin embargo, en este caso, las prestaciones para los tres tipos de distribuciones de error son similares.

	Distribución peatonal	Distribución vehicular	Distribución global
Peatonal	1886.9	631.5	1807.8
Vehicular	902.2	904.3	902.3

Tabla 13: Throughput medio en kbps.

Se puede concluir la utilidad del método como mecanismo de *link adaptation* en HSDPA. Nótese que para el caso vehicular se obtienen mejores prestaciones que empleando el bucle de corrección (véase la Tabla 6). Las buenas prestaciones obtenidas con la distribución global facilitan la implementación real al no precisarse la diferenciación de móviles según la velocidad.

Se plantea una mejora adicional consistente en utilizar un filtro de mediana sobre la serie de CQIs recibidos previamente a la obtención de las distribuciones. Con dicho filtro se pretende reducir el efecto del error de medida en la estimación. Tras la aplicación de dicho filtro y la simulación con las nuevas distribuciones de error se tendrían los resultados de la siguiente tabla:

	Distribución peatonal	Distribución vehicular	Distribución global
Peatonal	1922.1	780.5	1899.4
Vehicular	887.7	923.8	904.6

Tabla 14: Throughput medio en kbps.

A la luz de los resultados queda demostrada la utilidad de la aplicación del filtro ya que se obtiene una ganancia en prácticamente todos los casos y siempre para la distribución global. Nótese que el usuario peatonal, con la distribución peatonal puede incluso mejorar las prestaciones del bucle de control de la Tabla 5 y en cualquier caso, para la distribución global se obtienen resultados similares.

VI.5 *El problema de la carga baja.*

En esta subsección se aborda un problema diferente a toda la temática anterior. Dicho problema se presentó tras realizar un análisis de informes reales de comunicaciones con HSDPA.

Concretamente, se realizaron varias sesiones de navegación web. Los mensajes del chipset del móvil empleado fueron almacenados mediante una herramienta de análisis de tiempo real en una serie de archivos binarios. Posteriormente los archivos binarios generados fueron procesados mediante un software propiedad del Grupo de Comunicaciones Móviles obteniéndose una base de datos con los parámetros más significativos de la transmisión.

En general, las sesiones web se caracterizan por un tráfico muy variable. Pueden encontrarse periodos largos sin transmisión alguna y periodos activos en los que generalmente el tráfico es muy impulsivo, dándose una transmisión de bloques a ráfagas. Estas suposiciones sobre el tráfico web fueron claramente observadas con las medidas realizadas.

Se apreció que muchas veces, tras un periodo de inactividad, los primeros bloques en transmitirse llegaban erróneos. Además, se observaba una disminución en el valor de los CQIs enviados por el terminal al comenzar el periodo activo. Se dedujo que ambos fenómenos eran dos caras de la misma moneda. El problema radica en que en situación de inactividad y si ningún otro usuario recibiendo datos en la celda, la única potencia interferente recibida de la propia celda es la de los canales comunes que se transmiten continuamente. Mientras que cuando un usuario recibe información, le interfiere tanto la potencia de los canales comunes como la propia potencia enviada a él mismo. Este último fenómeno, común en los receptores WCDMA actuales, hace que el canal

que experimenta un usuario cambie mucho si pasa de inactivo a activo, en baja carga. De este modo, las estimaciones tomadas durante el tiempo de inactividad pueden no ser útiles y provocar fallos.

Es necesario cuantificar el error producido y corregirlo. Esta tarea es factible, pero no es trivial como se expone a continuación.

El SINR del canal percibido por un móvil a partir de la potencia del piloto, es :

$$SINR = \frac{P_{CPICH} + \Gamma}{P_{own}(1-\alpha) + P_{other} + Noise} = \frac{\varepsilon\, P_{BS}}{P_{own}(1-\alpha) + P_{other} + Noise} \tag{26}$$

, donde P_{CPICH} , P_{BS} , P_{own} , P_{other} y $Noise$ son la potencia recibida en el terminal móvil procedente del canal piloto, de la propia estación base en caso de radiar ésta su máxima potencia, de la propia estación base, del resto de celdas interferentes y del ruido térmico existente en el canal. Γ es la diferencia de potencia entre el canal piloto y la potencia asignada a una transmisión de datos. ε es la proporción de potencia de la estación base dedicada a una transmisión de datos. α es el factor de ortogonalidad ya comentado.

Con baja carga, la potencia recibida de la propia celda es una porción de la potencia máxima radiada por la celda:

$$SINR_{lowload} = \frac{\varepsilon\, P_{BS}}{\eta P_{BS}(1-\alpha) + P_{other} + Noise} \tag{27}$$

, donde η sería la proporción de potencia total de la estación base que se está radiando realmente.

Con carga alta, la potencia recibida de la propia celda es la potencia máxima radiada por la celda ($\eta = 1$):

$$SINR_{highload} = \frac{\varepsilon\, P_{BS}}{P_{BS}(1-\alpha) + P_{other} + Noise} \tag{28}$$

Si se dividen ambas expresiones, se tiene:

$$SINR_{highload} / SINR_{lowload} = \frac{\eta P_{BS}(1-\alpha) + P_{other} + Noise}{P_{BS}(1-\alpha) + P_{other} + Noise} \tag{29}$$

Definiendo el factor de geometría, G , como:

$$G = \frac{P_{BS}}{P_{other} + Noise} \tag{30}$$

se puede expresar la razón anterior como:

$$SINR_{highload} / SINR_{lowload} = \frac{\eta(1-\alpha) + 1/G}{(1-\alpha) + 1/G} = \frac{1 + \eta G(1-\alpha)}{1 + G(1-\alpha)} \tag{31}$$

Por tanto, la corrección debería depender de tres parámetros: η , α y G . El problema es que estos parámetros no tienen un valor fijo en el sistema así que la corrección a realizar en caso de ser constante no sería óptima.

A continuación se analizará un problema relacionado que se puede presentar. La estación base sabe cuándo se transmitieron datos en la celda (con su interferencia adicional) y cuándo no, así que sabe cuándo tiene que aplicar y cuándo no el factor de corrección. Pero, el canal descendente de datos no está alineado con el ascendente, de manera que el periodo de medida para la obtención del CQI coincide con dos subtramas descendentes. La segunda mitad de una y la primera mitad de la otra. Es decir, que cuando el terminal móvil mide el CQI puede que: mida en una subtrama en la que no hay nada de interferencia de datos, mida en una subtrama contaminada en una mitad por interferencia de datos o mida una subtrama totalmente contaminada por interferencia de datos. Esta diferente interferencia se podría ver como diferentes valores de η, valores que serían altos con interferencia alta y bajos con poca interferencia.

Con todo esto, la estrategia propuesta consiste en realizar una corrección diferente según la interferencia intracelular que el Nodo-B considera que se tuvo en la medida, obviando el resto de factores variables. La estrategia se puede resumir con la siguiente fórmula:

$$CQI_{corrected} = \begin{cases} CQI_{original} & si\ \eta = 1 \\ CQI_{original} - 3 & si\ \eta = \eta_{medium} \\ CQI_{original} - 5 & si\ \eta = \eta_{\min} \end{cases} \quad (32)$$

A continuación se comparan los resultados de varias simulaciones con usuarios web. Se consideran tanto usuarios peatonales como vehiculares, con y sin retransmisiones, con y sin corrección por baja carga.

La Tabla 15 muestra los porcentajes de elecciones de CQI optimistas, óptimas y conservativas. En cada celda de la tabla se muestran los porcentajes en el orden indicado y separados por barras. Por elección de CQI óptima se entiende que el CQI elegido para la transmisión de un bloque radio y el correspondiente a la SINR experimentada en la recepción de dicho bloque coinciden. La elección es conservativa si el SINR experimentado correspondiera a un CQI menor al elegido. Por el contrario, sería optimista si el SINR experimentado fuera menor que el asociado al CQI elegido. Al realizar la corrección disminuye la proporción de elecciones optimistas que se traducen en fallos de transmisión, mientras que aumenta la proporción de elecciones conservativas.

	Peatonal sin retransmisiones	Peatonal con retransmisiones	Vehicular sin retransmisiones	Vehicular con retransmisiones
Sin corrección	14.9 / 66.3 / 18.7	23.5 / 50.8 / 25.6	12.1 / 62.6 / 25.3	14.7 / 60.8 / 24.4
Con corrección	19.4 / 13.5 / 67.1	18.9 / 16.3 / 64.8	17.8 / 21.5 / 60.7	16.2 / 33.3 / 50.5

Tabla 15: Porcentajes de tipo de estimación del CQI (óptima / optimista / conservativa).

La Tabla 16 muestra los porcentajes de retransmisiones para los casos sin corrección y con corrección, tanto para usuarios peatonales como vehiculares. Se aprecia una clara mejora, en forma de reducción del número de retransmisiones necesarias, al añadir la corrección.

	Peatonal	Vehicular
Sin corrección	22.2	30.8
Con corrección	6.7	16.9

Tabla 16: Porcentaje de retransmisiones

La reducción en el número de retransmisiones aun así no es lo más importante. Podría darse el caso de que la corrección fuera demasiado conservativa, produciendo muy pocos fallos, pero siendo ineficiente. Esto es, puede que se requieran muchas transmisiones para mandar toda la información. Un buen indicador de la eficiencia de la transmisión web es el retardo normalizado de la misma. Se define como el tiempo real de transmisión dividido entre el número de bits transmitidos en ese tiempo. En la Tabla 17, se muestran las medianas en ms/kbit. La mejora al aplicar la corrección es evidente por la disminución del tiempo necesario para transmitir cada kilobit de datos. Por tanto, se puede concluir que la corrección propuesta es eficiente.

	Peatonal sin retransmisiones	Peatonal con retransmisiones	Vehicular sin retransmisiones	Vehicular con retransmisiones
Sin corrección	3.83	1.61	7.72	2.25
Con corrección	1.27	1.21	1.68	1.62

Tabla 17: Medianas del retardo normalizado en ms/kbit

VII. CONCLUSIONES.

La presente tesina ha consistido en un profundo estudio del *link adaptation* en HSDPA.

Por un lado se ha evaluado el procedimiento de adaptación al enlace más común en la bibliografía, concluyéndose su general ineficiencia para maximizar el *throughput* experimentado por un usuario.

Para mejorar las prestaciones del algoritmo de bucle de corrección del CQI se ha propuesto la operación conjunta de este proceso y de un prefiltrado o preprocesado de los CQIs. Se ha empleado en el estudio un preprocesado basado en la predicción de los CQIs con el método LMS previo filtrado de la serie de CQIs recibidos con un filtro promediador. La elección de éste procesado se debe a que ha demostrado unas buenas prestaciones en otros escenarios [17]. Los resultados obtenidos tras la aplicación de este método lo han validado como una opción a tener en cuenta para mejorar las prestaciones del *link adaptation* más común.

Por otro lado, a partir del análisis de un *link adaptation* genérico, se ha propuesto una estrategia de *link adaptation* que tiene en cuenta la estadística de la variación del CQI. Dicha estadística se recoge en una serie de funciones que relacionan cada CQI medido con la probabilidad de ocurrencia de ése u otros CQIs un tiempo después. En base a resultados, se ha demostrado que la estrategia propuesta es capaz de igualar o mejorar las prestaciones del mecanismo de *link adaptation* común.

Por último, se ha estudiado un problema del *link adaptation* de HSDPA que se produce en situaciones de carga baja. Se ha propuesto una corrección de los CQIs recibidos basada en la actividad del Nodo-B. El mecanismo propuesto ha sido evaluado mediante simulación comprobándose su utilidad.

AGRADECIMIENTOS

La presente tesina ha sido realizada en colaboración con el proyecto “Gestión cooperativa de redes de acceso radio heterogéneas mediante técnicas predictivas”, TEC2005-08211-C02, financiado por la Comisión Interministerial de Ciencia y Tecnología.

El autor de la tesina desarrolla su actividad investigadora en el marco del Programa de Formación de Profesorado Universitario (FPU).

REFERENCIAS

[1] H.Holma, A.Toskala, *HSDPA/HSUPA for UMTS,* John Wiley and Sons, 2006.

[2] E.Dahlman, S.Parkvall, J.Sköld and P.Beming, *3G Evolution: HSPA and LTE for Mobile Broadband*, Academic Press, 2007.

[3] H.Holma, A.Toskala, *WCDMA for UMTS: Radio Access for Third Generation Wireless Communications*, Third Edition, John Wiley and Sons, 2004.

[4] P.J.Ameigeiras, *Packet Scheduling and Quality of Service in HSDPA,* Ph.D. Thesis, Aalborg University (Denmark), October 2003.

[5] 3GPP Technical Specification 03.64, *Overall Description of the GPRS radio interface, Stage 2;* v. 8.2.0, 2004.

[6] P.J.Ameigeiras, J.,Wigard, P.N.Andersen, H.C.Damgaard, P.Mogensen, *Performance of Link Adaptation in GPRS Networks*, Proc. IEEE Vehicular Technology Conference Fall 2000, vol.2, pp. 492-499.

[7] O.Keseth, F.Gessler and M.Frodigh, *Algorithms for Link Adaptation in GPRS*, Proc. IEEE Vehicular Technology Conference 1999, vol.2, pp. 943-947, May 1999.

[8] 3GPP Technical Specification 45.008, *Radio subsystem link control,* v 7.1.0, 2005.

[9] S.Eriksson, A.Furuskär, M.Höök, S.Jäverbring, H.Olofsson and J.Sköld, *Comparison of Link Quality Control Strategies for Packet Data Services in EDGE*, Proc. IEEE Vehicular Technology Conference Spring 1999, Houston, pp 938-942, May 1999.

[10] 3GPP Technical Specification 25.214, *Physical Layer procedures (FDD),* v. 7.6.0., 2007.

[11] 3GPP Technical Specification 34.121-1, *User Equipment (UE) conformance specification;Radio transmission and reception (FDD); Part 1; Conformance specification*, v. 8.0.0., 2007.

[12] K.I.Pedersen, F.Frederiksen, T.E.Kolding, T.F.Lootsma and P.E. Mogensen, *Performance of High-Speed Downlink Packet Access in Coexistence With Dedicated Channels*, IEEE Transactions on Vehicular Technology, vol. 56, no. 3, May 2007.

[13] M.Nakamura, Y.Awad and S.Vadgama, *Adaptive Control of Link Adaptation for High Speed Downlink Packet Access (HSPDA) in W-CDMA*, Proc. WMPC, 2002, pp. 382-386.

[14] D.W.Parachynch and M.Yavuz, *A Method for Outer Loop Rate Control in High Data Rate Wíreles Networks*, Proc. IEEE Vehicular Technology Conference 2002, pp.1701-1705.

[15] J.Gozálvez, D.Martín-Sacristán, M.Lucas-Estañ, J.F.Monserrat, J.J.Gonzalez-Delicado, D.Gozálvez and M.Marhuenda, *SPHERE – A Simulation Platform for Heterogeneous Wireless Systems*, IEEE Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (Tridentcom), Orlando, 2007.

[16] J.F.Monserrat, D.Martín-Sacristán, D.Gozálvez, N.Cardona and J.Gozálvez, *SPHERE – A Simulation Tool for CRRM Investigations*, COST2100, Feb 2007.

[17] D.Martín-Sacristán, J.F.Monserrat, J.Gozálvez and N.Cardona, *Effect of Channel Quality Indicator Delay on HSDPA Performance*, Proc. IEEE Vehicular Technology Conference Spring 2007, pp. 804-808.

[18] D.Martín-Sacristán, J.F.Monserrat, D.Calabuig and N.Cardona, *HSDPA Link Adaptation Improvement Based on Node-B CQI Processing*, International Symposium on Wireless Communications Systems 07, Trondheim, Oct 2007.

[19] European Telecommunications Standards Institute, *Selection Procedures for the choice of radio transmission technologies for the UMTS*, TR 101 112 (UMTS 30.03), v. 3.2.0, April 1998.

[20] M.Gudmundson, *Correlation Model for Shadow Fading in Mobile Radio Systems*, Electronics Letters, vol. 27, no. 23, pp. 2145-2146, November 1991.

[21] H. Kim and Y. Han, *Enhanced Correlated Shadowing Generation in Channel Simulation*, IEEE Communications Letters, Vol. 6, pp 279-281, 2002.

[22] Lucio Ferreira et alii, *Definition of Reference Scenarios for the evaluation of radio resource allocation algorithms*, Department7 , Activity AP1, NEWCOM, June 2006

[23] K.I.Pedersen and P.E.Mogensen, *The Downlink Orthogonality Factor Influence on WCDMA System Performance*, Proc. IEEE Vehicular Technology Conference Fall 2002, vol.4, pp. 2061-2065.

HSDPA Link Adaptation Improvement Based on Node-B CQI Processing

David Martín-Sacristán, Jose F. Monserrat, Daniel Calabuig and Narcís Cardona

Mobile Communications Group, Polytechnic University of Valencia (UPV)

Valencia, Spain

{damargan,jomondel,dacaso,ncardona}@iteam.upv.es

***Abstract* — In this paper HSDPA link adaptation (LA) based on Channel Quality Indicator (CQI) reports is optimised. A pre-processing of the last received CQI reports is done before the execution of the LA algorithm in the Node-B in order to obtain more profitable channel quality estimations and hence improve the LA performance. Different types of processing techniques are presented and assessed, considering from the simplest sample averaging to some more elaborated predictive algorithms. Results demonstrate that a non negligible enhancement in the LA performance can be obtained if medium and high speed users are considered.**

I. Introduction

High Speed Downlink Packet Access (HSDPA) and its uplink counterpart HSUPA have been recently standardized in 3GPP to improve the performance of previous UMTS systems. HSPA standards aim at increasing the downlink packet data throughput while efficiently sharing the available radio resources. HSDPA achieves these objectives by means of adaptive modulation and coding (AMC), fast scheduling mechanisms (each TTI or Transmission Time Interval of 2 ms) and a Hybrid ARQ mechanism. A good survey of HSPA principles can be found in [1].

Link adaptation (LA) and scheduling are processes of paramount importance to optimise HSDPA system performance. However, they have not been standardised to propitiate competitiveness among vendors and therefore many investigations are done in these fields.

In HSDPA, the user equipment (UE) reports periodically to the Node-B its experienced downlink channel quality by means of the Channel Quality Indicator (CQI) [2]. Numerically CQIs are integers extending from 1 to 30, increasing its value when channel quality augments.

Usually LA is based on these CQI reports in such a way that the resources allocated to one user are at most or even exactly the corresponding to its last reported CQI (or a slightly modified version which takes into account the actual availability of power, codes, etc.). Each CQI can be translated into a combination of transmission parameters since [2] establishes a relation between each CQI and a concrete value of the transport block size (TBS), number of simultaneous channelisation codes, modulation and code rate.

However, some time passes since the CQI is measured by the user until this information is employed by de Node B to perform a downlink transmission. In particular, if the minimum CQI reporting period (2ms) is considered, this delay is around 7ms. Due to the fast radio channel variations and the existence of this non negligible delay, the last CQI report sent by the user may not be reliable to carry out the LA, most of all when the user velocity is high. The effect of this inaccuracy on the system performance has been studied in several papers as [3] and [4], showing that this imprecision is critical for the system performance. In consequence, some additional mechanisms are needed to correct the CQI inaccuracy or to extract useful information from the inaccurate CQI reports. A method to correct the reported CQIs employing the ACK-NACK ratio is proposed in [1] and [5], while [6] presents a strategy which reduces the employed CQI, thereby increasing data protection, when the time elapsed since the last CQI report increases. All these algorithms prevent the system from making mistakes in LA but are unable to follow the channel state and therefore do not maximise the system capacity use.

The main objective of this paper is to optimise link adaptation via an estimation of the channel quality experimented by the UE in the next downlink transmission (in the form of a CQI) based on the knowledge of the past CQIs. With this aim, different types of pre-processing of the CQI reports received by the Node-B are presented and assessed.

The rest of the paper is organised as follows. In Section II the employed system level simulation tool is described. In Section III different techniques of CQI pre-processing are presented, comparing its capacity to predict the future channel state. The effect of the pre-processing techniques on HSDPA LA performance is evaluated in Section IV. Finally, the main conclusions are drawn.

II. Simulation Environment

Simulation in this paper has been conducted with the emulator presented in [4], which is part of the SPHERE simulation platform described in [7]. In Section III the emulator has been employed to obtain CQI report traces whereas in Section IV it has been used to evaluate the system level performance of CQI processing. Table I summarises the most important parameters of the simulated scenarios.

TABLE I Most Important Simulation Parameters

Parameter	Value
Load	1 user
Simulation time	1800 s (vehicular scenarios) , 7200s (pedestrian scenarios)
Cell radius	1 Km
Number of cells	7 (central cell and one tier of interferers)

UE Speed	50 km/h (vehicular) or 3 km/h (pedestrian)
Node-B Tx Power	43 dBm
CPICH Power	33 dBm
Orthogonality factor	0.8
Path Loss model	$L(dB) = 128.1 + 37.6 \log_{10} d(Km)$
Shadowing model	Lognormal distribution, 8dB std dev
Fast fading model	$f_{\psi}(\psi) = \frac{27}{2}\psi^2 e^{-3\psi}$
UE category	10
Scheduling algorithm	Round Robin
HARQ SAW processes	6
Retransmissions	Not considered

III. CQI PROCESSING

In this section the variability of the CQI reports is studied. Next, some processing strategies are presented and its ability to perform effective CQI estimations is compared.

A. CQI variability analysis

In HSDPA the UE performs periodical CPICH *CINR* (Carrier to Interference plus Noise Ratio) measurements, mapping each calculated *CINR* to a concrete CQI whose transmission parameters would ensure a *BLER* (Block Error Rate) equals to 10%. Therefore, the CQI variability is directly related to the *CINR* variability, the only difference is that the CQI is like a quantization of the *CINR* with 1 dB steps.

The *CINR* is usually calculated in HSDPA with the following equation:

$$CINR = \frac{P_{CPICH} \cdot \Gamma \cdot L \cdot S \cdot F}{(1-\alpha) P_{total} L \cdot S \cdot F + \sum_{i=1}^{I} P_{total_i} L_i S_i F_i + N} \qquad (1)$$

where the numerator contains the useful power received in the UE antenna and the denominator consists of three terms: the first one represents intra-cell interference, the second one represents inter-cell interference and the latter, *N,* represents the thermal noise. P_{CPICH} is the CPICH power, P_{total} is the total power transmitted by the serving Node-B, P_{total_i} is the total power transmitted by each one of the *I* interfering Node-B. *L*, *S* and *F* represent the path losses, shadowing and fast fading respectively. L_i, S_i and F_i are the path losses, shadowing and fast fading of the link between the user and the *i*-th interfering Node-B. Γ is the power offset between the CPICH power and the HS-PDSCH power. Finally, α is the orthogonality factor ranging from 0 to 1, meaning $\alpha = 1$ a perfect orthogonality.

The following factors explain the *CINR* variability:

1) L: Its dynamic margin depends on the cell radius and its variation rate depends on the mobile speed.

2) *S:* Its dynamic margin is fixed in the simulations (8dB standard deviation) and its variation rate is clearly related to the mobile speed (*S* is decorrelated after 1s for vehicular users at 50 km/h and 24s for pedestrian users at 3 km/h).

3) F: Its dynamic margin is fixed in the simulations (2.7dB standard deviation) but its variability depends on the speed (coherence time equals 2ms for 50 km/h and 32 ms for 3km/h).

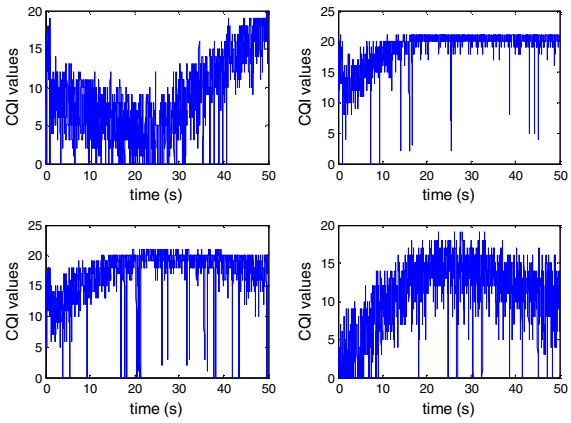

Fig. 1 CQI traces for pedestrian users during 50 seconds

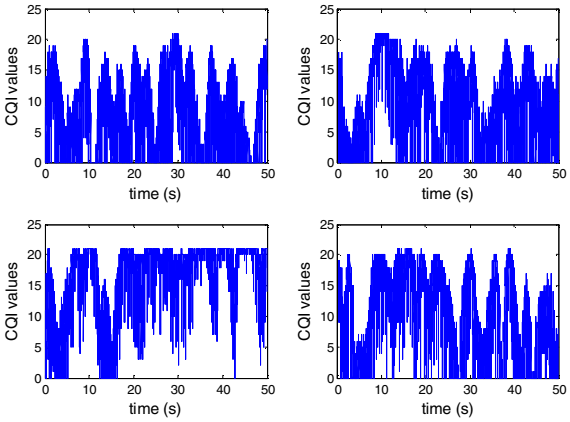

Fig. 2 CQI traces for vehicular users during 50 seconds

4) α *:* A perfect characterization of the *CINR* would require a variable orthogonality factor [8]. For the sake of simplicity, in this paper a fixed value of 0.8 is used.

5) P_{total} *:* The power transmitted by the Node-B is not constant. In low load scenarios, the transmitted power can change in a non negligible way from one TTI to another, introducing impulsive interference and hence short term CINR and CQI variations. However, in high load scenarios, this variation is small and a high almost-constant interference is expected to be received. To conduct this assessment a high loaded network has been considered.

To resume, in the short term fast fading variations and transmitting power fluctuations are the most prominent factors. In the mid term, shadowing effect is quite significant whereas the average fast fading and power fluctuations effect is zero. Finally, in the long term, the path losses changes are the only important factor. In this paper the short term variation is the most important one since the main objective is to perform short term CQI estimations.

Next, some CQI traces are presented to show CQI variability. Two figures have been plotted analysing different scenarios. Fig1 considers pedestrian users moving at 3 km/h and Fig2 considers vehicular users moving at 50 km/h. Pedestrian and vehicular traces are similar but the main difference is the faster variation in the CQI vehicular traces.

These traces have been obtained from a saturated scenario, what means that users always have information pending for transmission. For this scenario and having into account the considered orthogonality factor, the ratio between useful power and own cell interference determines a maximum CQI value, which in this case is 21 as corroborated in the figures.

B. Future CQI estimation strategies

According to the methodology of processing, three different classes of estimation strategies have been analysed. The first one is the simplest scheme which takes the last reported CQI as the future CQI estimation (mode 1). The second class comprises different strategies based on the filtering of the last *n* reported CQI, being this filtering linear (modes 2, 3, 5, 6) or non linear (mode 4). The third class is composed by predictive schemes which perform an adaptive filtering of the last reported CQIs to obtain a prediction of the future CQI (modes 7, 8).

Next, all the processing modes are explained:

1) *No processing* (No proc): the basic strategy in which the last reported CQI is directly the future estimation.

2) Exact averaging (Averaging):the estimation is obtained averaging the last *n* reports.

3) *Smoothed averaging* (Smoothing): the estimation is obtained smoothing the reported trace of CQIs with the next formula:

$$\widehat{CQI}(k) = \left(1 - 1/T\right)\widehat{CQI}(k-1) + \left(1/T\right)CQI(k), \qquad (2)$$

where $CQI(k)$ is the last reported CQI, $\widehat{CQI}(k-1)$ is the old CQI estimation, $\widehat{CQI}(k)$ is the new CQI estimation and T is the equivalent averaging period in TTIs.

4) Median CQI (Median): the estimation is obtained as the median of the last *n* reports.

5) *Weighted conservative filtering* (W.cons.): the estimation is obtained as the weighted average of the last *n* reports, being the weighting function linear, giving more weight to the lower CQI reports

6) Weighted smart filtering (W.smart): the estimation is obtained as the weighted average of the last *n* reports, in this case giving more weight to the newest CQI reports. Again the weighting function is linear.

7) LMS prediction (LMS): the estimation is the prediction performed by the LMS method [9] with one coefficient and a step size of 2e-4.

8) *LMS average prediction* (LMS avg.): in this case the reference signal is the trace of the exact CQI averaging.

C. Future CQI estimation results

In the remaining of this section the performances of all the processing methods are compared. The minimum CQI reporting period of 2ms has been considered since it is a common used value. The effect of the number of samples used in the averaging has been assessed too.

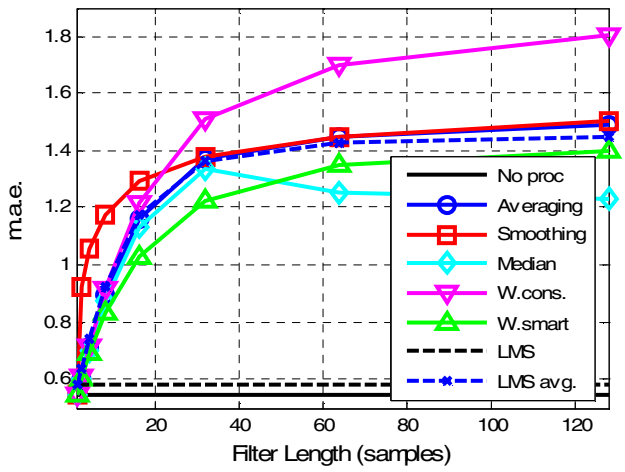

Fig. 3 Mean Absolute Error of the CQI estimation modes for different filter lengths. Pedestrian users

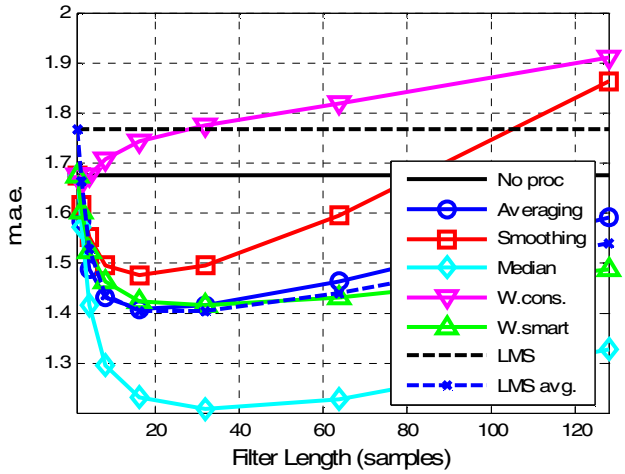

Fig. 4 Mean Absolute Error of the CQI estimation modes for different filter lengths. Vehicular users

Pedestrian users moving at 3 km/h and vehicular users moving at 50 km/h have been studied.

First of all, the mean absolute error (m.a.e.) of the estimations is shown in Fig 3 and Fig 4.

In the pedestrian environment (Fig. 3) the best results are achieved by the basic method. On the other hand, in the vehicular environment (Fig 4) this mode can be highly improved thanks to the use of processing techniques. The optimum filter length depends on the processing technique but is around 16-32 samples. This divergence between vehicular and pedestrian is explained by the different fast fading variability in each case. In the pedestrian scenario the channel does not practically change in 7 ms and hence the basic mode behaves well. Besides, all techniques based on averaging introduce an artificial longer delay in the channel tracking that can not improve the basic method. In a vehicular scenario channel changes faster and the basic mode is not able to track these variations. It is necessary to perform some kind of processing to obtain a better estimation of the mean future channel quality, which results in a better m.a.e.

Focusing on the vehicular scenario, one interesting observation is that exact averaging behaves better than smoothed averaging. In the latter, impulsive variations of the CQI affect the CQI estimation for a longer time. Regarding the non linear filtering methods, the median CQI scheme presents the best performances as compared with the other algorithms since this scheme removes the effect of deep fades in the CQI due to the fast fading, avoiding the pernicious effect of these outliers. The weighted conservative filtering performs conservative choices, not accurate choices, and therefore presents high m.a.e. values. The weighted smart scheme presents good results, similar to the averaging scheme due to its inherent ability to track the channel quality.

The predictive schemes behave similar than their non predictive counterparts. The LMS behaves like the basic method and the LMS applied to filtered traces behaves similar to the exact averaging. The simple LMS prediction performance is lower than it could be expected by the reader due to the low predictability of the series under study.

However, although the m.a.e. provides an interesting information about the ability of the different processing techniques to perform accurate estimations, it is not the only important factor to be considered in the LA optimisation. The success of the CQI estimation has been also analysed. A CQI is optimally estimated if the CQI selected to transmit and the CQI experienced in the reception are equal. An optimistic CQI estimation occurs when the CQI used in the transmission is higher than the experienced it the reception and therefore the information is not properly received. Finally, a conservative CQI estimation is done when the experienced CQI is lower.

Next figures show the ratios of optimal, optimistic and conservative CQI choices for the considered algorithms. The same filter length of 64 samples has been employed for the pedestrian and vehicular scenarios. As expected, filtering increases the conservative CQI choice. This effect is more visible with the most conservative algorithms. Predictive schemes show a similar effect but less pronounced.

As it can be observed in Fig. 5, in a pedestrian scenario the basic mode of no processing behaves really well. All the other methods produce a decrease in the optimal choice ratio and an increase in the optimistic and/or conservative ratios. In the vehicular environment the basic mode presents a poorer behaviour while other methods obtain more accuracy, like in the median case, or increase the ratio of conservative choices.

IV. System Level Simulations

The results of Section II measure the accuracy of the CQI estimations and how the different algorithms can change the ratio of optimistic, optimal and pessimistic CQI selections as compared with the basic mode. In this section the effect of the CQI processing technique on the LA performance is evaluated in a simulated high-loaded network.

The simulation considers a single user per cell since in such a scenario the results are independent of the scheduling algorithm. The high load has been emulated forcing the users to have always data pending for transmission in the Node-B. Pedestrian and vehicular users were simulated.

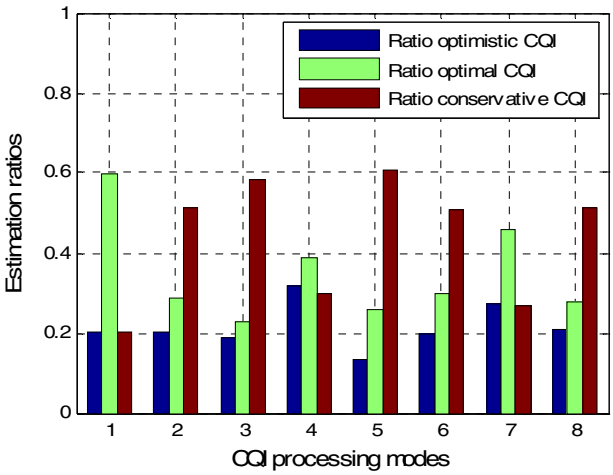

Fig. 5 Ratio of optimistic, optimal and conservative CQI selection for each processing mode. Pedestrian users.

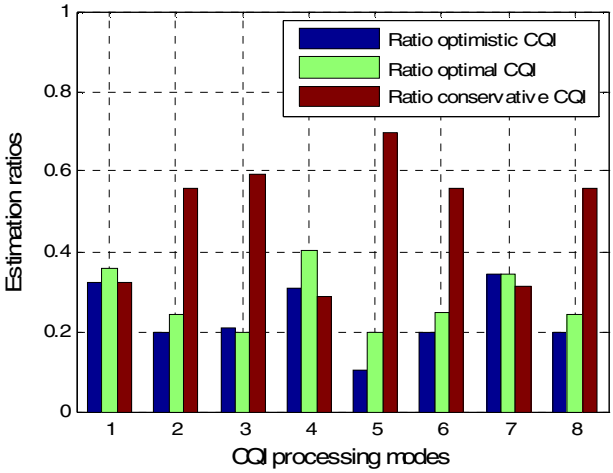

Fig. 6 Ratio of optimistic, optimal and conservative CQI selection for each processing mode. Vehicular users.

Two parameters have been evaluated, namely throughput, i.e. bits transmitted per second, and block error rate (BLER), defined as the ratio between the number of blocks correctly received and the total number of transmitted blocks.

In Section II it was concluded that the processing of the CQI reports can enhance the basic channel quality estimation but only in a vehicular scenario. The results obtained in the system level simulations corroborate this idea. As shown in Fig. 7 and Fig. 8, the maximum throughput for the pedestrian case is achieved with the basic method whereas in the vehicular scenario some gain can be obtained when processing is done.

With respect to the vehicular case, five of the proposed techniques obtain better results than the basic mode. LMS with filtered reference signal provides the best results (5.8% improvement) but the use of exact averaging and median modes entails a similar gain, as can be observed in Table II. Therefore, the gain is due basically to the averaging but not to the prediction.

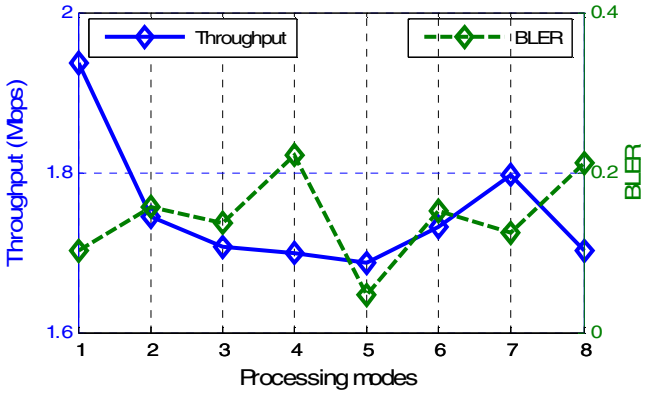

Fig. 7 Throughput and BLER in a pedestrian high-loaded scenario.

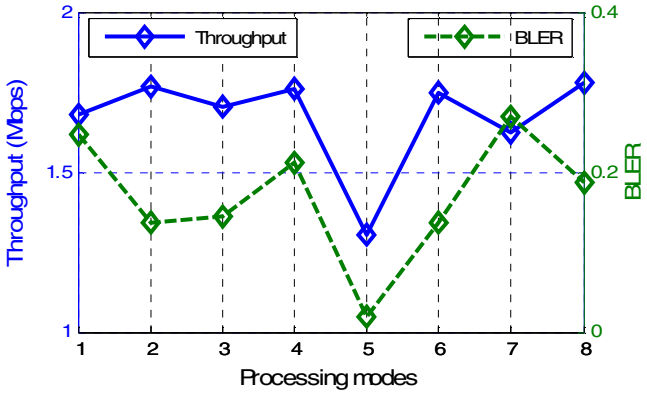

Fig. 8 Throughput and BLER in a vehicular high-loaded scenario.

TABLE II THROUGHPUT IMPROVEMENT IN THE VEHICULAR SCENARIO

Mode	2	3	4	5
Improvement (%)	5.10	1.33	4.89	-22.40
Mode	6	7	8	
Improvement (%)	4.26	-3.22	5.80	

According to the Fig. 8, and taking into account the information of Figure 6, it can be concluded that there are two ways of improving the system performance. The first one is to improve the optimal choice ratio, which is what happened with the median method. In this case, both the number of incorrectly received blocks (directly related to the optimistic choices) and the number of transmissions where the link resources are not fully profited (directly related to the conservative choices) is reduced. Therefore, the optimality as was considered in the CQI selection is highly related to throughput maximization. On the other hand, methods 2, 3, 6 and 8 accomplish the objective of improving the throughput via a more conservative CQI estimation than the basic mode. Although they reduce the optimal choice ratio and increase the conservative choices they lessen substantially the optimistic choices and hence the BLER. This decreasing in the BLER justifies the throughput improvement since fewer blocks are lost. Method 5, the more conservative, reduces the BLER but is too much conservative and its choices do not take advantage of the channel quality as other algorithms do. Finally, method 7 is unable to improve the system performance since it augments the BLER by increasing the optimistic choice ratio.

Two important matters arise from now. The first is the determination of the threshold speed in order to apply or not the CQI processing. According to simulations this threshold speed is around 16 km/h for the simulated scenarios. The second one is the evaluation of how CQI measurement errors affect the performance of the CQI processing. Again, according to simulations, assuming a 1dB standard deviation lognormal error in the *CINR* CPICH measurements it has been proven a LA performance enhancement higher than the one obtained in this paper, which was around 15%.

V. CONCLUSIONS

In this paper it has been evaluated the effect of applying a CQI pre-processing before performing LA in the HSDPA Node-B. As explained, usually the last reported CQI is the only input to the LA algorithm but this strategy can be suboptimal when the radio channel quality changes quickly. It has been demonstrated that when the UE speed is high some processing strategies can provide a better channel estimation than the last received CQI. This better channel estimation results in an increasing of the system throughput up to the 5.8% and a reduction of the BLER. On the contrary, when the UE speed is low, the basic mode outperforms other processing techniques, which is justified by the high channel coherence time. Therefore, the implementation of a CQI processing technique in a real system will require a previous knowledge of the UE speed in order to perform or not the processing technique.

ACKNOWLEDGMENT

This work has been partially funded by CICYT (Spanish National Science and Technology Council) and the FEDER program of the European Commission under the Project TEC2005-08211-C02.

REFERENCES

[1] H. Holma, A. Toskala, *HSDPA/HSUPA for UMTS: High Speed Radio Access for Mobile Communications*, John Wiley and Sons, 2006.
[2] 3GPP, Physical layer procedures (FDD), 3GPP TS25.214.
[3] T.E. Kolding, "Link and system performance aspects of proportional fair scheduling in WCDMA/HSDPA," IEEE Vehicular Technology Conference, vol 3, pp 1717-1722, October 2003.
[4] D. Martín-Sacristán, J.F. Monserrat, J. Gozálvez and N. Cardona, "Effect of Channel Quality Indicator delay on HSDPA performance," IEEE Vehicular Technology Conference, April 2007.
[5] M. Nakamura, Y. Awad, and S. Vadgama, "Adaptive control of link adaptation for high speed downlink packet access (HSDPA) in WCDMA," International Symposium on Wireless Personal Multimedia Communications, Oct. 2002.
[6] A. Muller, Tao Chen, "Improving HSDPA link adaptation by considering the age of channel quality feedback information," IEEE Vehicular Technology Conference, vol 3, pp 1643-1647, Sept. 2005.
[7] J. Gozálvez, D.Martín-Sacristán et alt., "SPHERE – A Simulation Platform for Heterogeneous Wireless Networks," IEEE TridentCom, May 2007.
[8] S. Burger, H. Buddendick, G. Wolfle, P.Wertz, "Location dependent CDMA orthogonality in system level simulations," IEEE Vehicular Technology Conference, vol 1, pp 419-423, June 2005.
[9] S.Haykin, *Adaptive Filter Theory*, 4th edition, Prentice Hall, 2001.

Effect of Channel-Quality Indicator Delay on HSDPA Performance

David Martín-Sacristán[1], Jose F. Monserrat[1], Javier Gozálvez[2] and Narcís Cardona[1]
[1] *Polytechnic University of Valencia (UPV) – Mobile Communications Group, Spain*
[2] *University Miguel Hernández – Signal Theory and Communications Division, Spain*

damargan@teleco.upv.es, jomondel@dcom.upv.es, j.gozalvez@umh.es, ncardona@dcom.upv.es

Abstract- **This paper evaluates the effect of the channel estimation inaccuracy on the performance of an HSDPA system. This study provides some results from system level simulations that have been conducted over a very complete dynamic simulator which models an HSDPA system full compliant with specifications. This emulator allows performing multi-user transmission and link adaptation with a limited modulation and coding scheme (MCS) selection based on the channel quality indicator (CQI) modes. Many factors such as the user equipment (UE) speed, the employed scheduling algorithm or the traffic load have been considered in the assessment. Moreover the intrinsic constraints of a WCDMA system like HSDPA have been also taken into account, i.e. the maximum number of channelization codes and the maximum transmitted power have been modelled jointly with a complete and dynamic interference characterization.**

Keywords- **HSDPA, CQI, Scheduling, Link Adaptation.**

I. Introduction

In order to allow high data rate transmission in Wideband Code Division Multiple Access (WCDMA) third generation mobile networks, High Speed Downlink Packet Access (HSDPA) and its uplink counterpart HSUPA have been recently standardized in 3GPP. HSDPA is already in a very mature state, and first commercial systems are being rolled out in the course of this year. HSDPA achieves high data rates of up to 14 Mbps by means of adaptive modulation and coding, fast scheduling mechanisms (each TTI or Transmission Time Interval of 2 ms) and a powerful Hybrid ARQ mechanism. Link adaptation (LA) is a process of paramount importance to optimise system functioning and therefore user equipment reports channel state either cyclically or in a triggered-based manner by means of the Channel Quality Indicator (CQI).

Several simulation results for HSDPA have been published in conferences and journal papers during last years. A good survey of HSDPA principles and performance simulations can be found in [1]. In [2] a closed formulation was proposed for the calculation of CQI whereas in [3] and [4] different reporting schemes were analyzed evaluating the effect produced by increasing report periods. Several scheduling algorithms have also been evaluated (see e.g. [5]) studying how the algorithm behaves for a constant bit rate traffic pattern.

This paper assesses the intrinsic delay of the Channel Quality Indicator (CQI) reporting process and its effect on the HSDPA system performance. As it is demonstrated in this paper, this delay has a severe impact in the efficiency of the link adaptation process and also affects the performance of the scheduling mechanisms since the upper limit of the resources allocated to each user depends on this report. This paper accomplishes this analysis in a mixed traffic scenario which consists of both best effort (web) and real time traffic (H.263 video telephony). A proper modelling of the HSDPA system constraints is implemented and a complete characterization of interferences is proposed. That is, to the best authors' knowledge, the singularity of this research as compared with for example [6] or [7]. The former considers an ideal traffic source and simplifies the interference variability whereas the latter uses a very simple model of the channel estimation inaccuracy.

The presented results prove that not only the mobile speed determines to what extent the CQI reporting delay has an effect on the system performance but also the scheduling algorithm and the system load among other factors.

This paper is organised as follows. First, the LA mechanism of HSDPA is briefly cleared up highlighting the time elapsed since the user sends its CQI report till the Node-B transmits according to this information. In Section III, all the fast packet scheduling algorithms compared in this paper are described. Next, in Section IV the evaluation environment is presented. Finally, the results of different simulations are presented and discussed and the most important conclusions are summarised.

II. Link adaptation Based on User Reports

In HSDPA the user equipment (UE) is responsible for reporting periodically to the Node-B the downlink channel quality. This channel information is numerically represented by means of the CQI, whose definition is explained in detail in [8]. Numerically CQI extends from 1 to 30, increasing its value when channel quality augments. Each CQI can be translated into a combination of transmission parameters since [8] establishes a relation among each CQI and a concrete value of the transport block size (TBS), number of simultaneous channelisation codes, modulation and code rate. These parameters were specifically chosen to configure a 1dB granularity in the carrier to interference ratio (CIR) among consecutive CQIs for a Block Error Rate (BLER) level of 10%.

As the channel state information is provided by the UE to the Node-B there is a non-negligible delay between channel estimation and the reception of this information in the Node-B. Moreover, some time pass since the Node-B receives the CQI until it uses this information in the LA and scheduling mechanisms. Besides, at least 2 slots more pass until the UE begins to receive the data scheduled in the Node-B due to the constant delay between the HS-SCCH channel which signals the start of data transmission and the HS-PDSCH which handles the data. The whole delay is around 6 ms, that is to say, three TTIs.

All of the previous delays are inherent in the HSDPA functioning and hence are not controllable. This paper deals with the only delay which can be controlled by the system, i.e. the reporting period delay.

In general, the higher the reporting delay is the higher inaccurate the channel estimation is. This inaccuracy depends on how the channel and therefore the CIR level vary in the time domain. This paper assesses the effect of this inaccuracy on the system performance depending on different factors as for example the scheduling algorithm employed.

III. Fast Packet Scheduling

Fast packet scheduling performed in Node B is one of the main features of HSDPA. Its implementation is not specified and investigation in this field can provide a great differentiation among different HSDPA systems.

In each TTI the scheduler makes a decision about to which UEs the Node-B will transmit in the next TTI and the characteristics of this transmission (TBS, number of channelization codes, code rate and power).

The scheduling decision is made taking into account a list of candidate UEs which have data to be delivered in the Node-B buffer. The scheduler can differentiate if these data are new or are waiting for retransmissions, giving a highest priority to retransmissions. A maximum number of code multiplexed users is fixed 'a priori' in the simulations, setting a limit to the number of multiplexed UEs in each TTI.

Category 10 UEs have been considered, which are the most flexible equipments, and only 30 modes of transmission, those corresponding to the 30 modes defined by the CQI table for this category, are employed. In spite of this restriction, the number of modes is enough to correctly consider the flexibility of HSDPA. Moreover, it has been demonstrated that such a resource allocation model can be more robust to the channel estimation inaccuracy than a more flexible one when the channel estimation inaccuracy is within some limits [7].

In this paper two kinds of resource allocating schemes are analyzed. The former does not guarantee any quality of service (QoS) since it employs scheduling algorithms that are not QoS aware. The latter is based on separating different types of services what can preserve real time services from best effort services.

Within the first group some classic scheduling algorithms such as Round Robin, Max-CIR and Proportional Fair are evaluated. Depending on the algorithm, a different prioritising scheme is established among the scheduling candidates. Round Robin scheduler prioritises users with an oldest last serving time, MaxCIR gives priority to the users with the highest last reported CQI (the best channel quality) while Proportional Fair prioritises users with the highest fairness factor (Ff) in the current time interval k defined as:

$$Ff_i(k) = R_i(CQI(k)) / \overline{R_i(k)} \quad (1)$$

where $R_i(CQI(k))$ is the maximum data rate that UE i can transmit provided the last reported CQI and $\overline{R_i(k)}$ is the mean data rate of the UE i in the time k, which is updated every TTI the UE has data waiting for transmission in the buffer according to the next formula:

$$\overline{R_i(k)} = \left(1 - 1/T\right)\overline{R_i(k-1)} + \left(1/T\right)R_i(k) \quad (2)$$

where T is the number of TTI considered in the averaging period, $\overline{R_i(k)}$ is the updated mean rate, $\overline{R_i(k-1)}$ is the old mean rate and $R_i(k)$ is the last instantaneous data rate.

Once the candidate list has been ordered the scheduler allocates resources. Generally the retransmissions are first served and afterwards the new transmissions if there are codes and power left. The scheduler processes a lot of information to make a decision as for instance the channel estimation reported by each user, the buffer size, the power available in the Node-B for transmission and, what is more important, the difference between this quantity and what the UE considered in the channel estimation.

Once a user has been served, the process is repeated with the next UE in the prioritised list until there is no more power or codes left. The objective of this scheme is to allocate the minimum power and the most efficient combination of transport block size and code rate to ensure a block error rate (BLER) of around 10% if possible.

All the afore-mentioned scheduling algorithms belong to the first group and provide a more or less good performance when handling best effort traffic but, in order to fulfill the QoS of real time (RT) users, other strategies must be considered. Therefore, another kind of resource allocating schemes is needed. The simplest option is to serve first the RT users and later the best effort traffic. This scheme is able to maintain the QoS for RT users if a call admission control (CAC) mechanism prevents congestion in the cell. This simple differentiation of services provides great results when considering a mixed scenario with RT and best effort users as compared with the case in which there is not differentiation. This aspect is clearly appreciated in the results presented in section V.

IV. Evaluation environment

To conduct this investigation an evolved version of the emulator presented in [9] has been employed, emulating HSDPA packet data transmission.

The simulator models a multi-tier macrocellular architecture with one hexagonal central cell with radius of 2800 or 1000 meters and 3 additional cell tiers. Simulations are only

conducted in downlink since HSDPA is specific of this direction. Users are on the move within the cell radius.

The available number of codes has also been carefully taken into account, and the same assumptions as in [10] have been made, assuming a value of 1 for the soft handover overhead. The maximum number of HS-SSCH codes has been set to 4, therefore code multiplexing of 4 users per TTI is allowed.

The HARQ mechanism of HSDPA has been implemented in a very realistic way. A stop and wait (SAW) protocol with 6 parallel processes has been considered in order to control the transmission to each UE.

The available power is modelled assuming a power consumption for the non data channels of HSDPA according to that used in [10]. The total Node-B power is 43 dBm.

Channel modelling comprises path loss, shadowing and also fast fading. Fast fading modelling is quite important when considering technologies, such as HSDPA, that base their radio operation on link adaptation techniques.

Intra-cell interference on a CDMA system is modelled by means of an orthogonality factor, which is usually denoted as α. In absence of multi-path fading, the codes are perfectly orthogonal and therefore $\alpha = 1$. In the worst case $\alpha = 0$, meaning that orthogonality is entirely destroyed. Typical values of α are between 0.4 and 0.9. In this research a value of 0.8 has been chosen. Thus, the HSDPA carrier to interference (*CIR*) level can be expressed as follows:

$$CIR_{HSDPA} = \frac{\frac{P_i}{L_P^{ii} \cdot L_S^{ii}} \psi_i}{\frac{(P_{Ti} - P) \cdot (1-\alpha)}{L_P^{ii} \cdot L_S^{ii}} \psi_i + \sum_{j \in \Omega} \frac{P_{Tj}}{L_P^{ij} \cdot L_S^{ij}} \psi_j + N_0 \cdot W} \quad (3)$$

where P_i is the addition of the power transmitted in all the channels allocated by the reference cell to the user of interest, P_{Ti} is the total power transmitted by the reference cell, Ω is the set of cells interfering the user and P_{Tj} is the total power transmitted by these interferers. In this expression, the parameters P_{Ti} and P_{Tj} also include the base station power reserved for other channels different from the HSDPA High Speed Downlink Shared Channel (HS-DSCH).

The simulation tool models all the interfering paths from each interfering base station to each UE. Other investigations consider a fixed level of interference coming from other cells or a fixed ratio between other cells interference and own cell interference. The more realistic approach implemented in this work allows a more accurate modelling of the interference variability and therefore of the channel quality variability. This accurate modelling is of paramount importance to assess the system performance most of all when radio access technologies based on link adaptation, as for instance HSDPA, are under consideration.

In the simulations presented in this paper several look-up tables (LUT) CIR vs BLER have been employed, one for each CQI value. These LUT are obtained from the European Network of Excellence NEWCOM [11].

In these LUTs it is also included the effect of the HARQ retransmissions with chase combining. In order to decide if a single block is correctly decoded or not the simulator computes the experienced CIR of this block in each slot of a TTI. After completing the transmission of a whole transport block, the three associated CIR values are averaged and a single CIR_{avg} value is obtained, which represents the quality experienced by the transport block. A LUT is employed to map the CIR value in a BLER value and to decide whether a block is correctly received. When a transport block is received in error, it is not discarded but stored in the receiver buffer and combined with retransmissions according to a specific method. The employed simulator uses the Chase Combining (CC) scheme in which retransmitted blocks are identical to that of the first transmission.

In the simulator web traffic has been modelled as a best effort traffic source. The web browsing service has been modelled as in [12]. It follows an ON/OFF pattern and a rate of around 55 kbps per user is expected due to the chosen parameters in [12]averaging over a long period of time.

Real-time services have also been included in the simulations through the emulation of real-time H.263 video transmissions following the model presented in [13].

V. System Performance

A meticulous simulation study has been carried out to assess the joint effect of CQI delay and processing and scheduling. Moreover different scenarios have been considered varying user speed and traffic patterns. The user speed is 3 km/h for the pedestrian users and 50 km/h for the vehicular users.

A. Performance with Saturated Traffic Sources

Initially, to determine the maximum cell capacity of the HSDPA system, users are simulated considering their traffic sources saturated. Each user has always 80 kbits pending for transmission in the serving Node-B buffer.

Regarding the cell throughput, the MaxCIR algorithm should achieve the highest performance since it allocates more resources to the users with the best channel quality, i.e. users with the highest available data rate. This fact can be observed in Table 1 and Table 2, which show the mean cell throughput for the MaxCIR, RR and PF algorithms. The cell throughput is defined as the total amount of bits correctly received per second in a cell. Both tables summarise the results obtained after 1800 seconds of system emulation with 15 users randomly distributed in the cell. Clearly the MaxCIR algorithm outperforms the other algorithms while the RR scheduler obtains the worst results as expected.

In addition, the relation between the CQI reporting period and the cell throughput is clarified in Table 1. It can be appreciated a reduction in the cell throughput when the reporting period increases. This effect is due to the fact that the CIR level changes dynamically due to the user movement, shadowing and fast fading and additionally due to the non negligible interference variability. Given that the reporting period increment is near milliseconds and the interference variability is highly reduced in a saturated scenario, the fast fading changeability is the most important factor which justifies the reduction in the cell throughput. This difference is higher in the pedestrian scenario than in the vehicular one. For

example, the difference in throughput is between 7% and 14% for the pedestrian case and between 0.2% and 3% for the vehicular case. This effect can be explained regarding the channel coherence time. In case of pedestrian users, fast fading is uncorrelated after 32ms whereas for vehicular users this time is reduced down to 2 ms. For this reason, in a vehicular scenario the reporting period increment has not a relevant effect on the system performance since already with 2ms the fast fading is uncorrelated and only the rest of factors, with slower variability, affect the inaccuracy of the channel estimation.

TABLE I. CELL THROUGHPUT IN MBPS FOR 1000M CELL RADIUS

	pedestrian			vehicular		
CQI reporting period (ms)	MCIR	PF	RR	MCIR	PF	RR
2	8.1466	4.6822	4.2911	7.5443	4.7024	4.1867
256	7.6087	4.1181	3.6732	7.5286	4.5412	4.1783

B. Performance with Best Effort traffic.

In the next scenario only web browsing users are considered. The study is focused in the effect of the CQI reporting period depending on the number of users, the scheduling algorithm and the user speed.

Table 2 collects some significant results in terms of the normalized delay experienced in the transmission of each web object when the MaxCIR algorithm is used. The normalized delay is defined as the delay in milliseconds required for transmitting one kbit of information. The mean value and the 95th percentile are shown in the Table. It can be observed that the higher the reporting period the higher the normalized delay, that is to say, the system performance is deteriorated since a higher normalized delay means more time to transmit the same data. Besides, increasing the number of users entails a higher normalized delay what is quite obvious in interference-limited systems as HSDPA. Moreover, these results reinforce the idea stated before, that is, in case of a vehicular scenario the degradation experienced with increasing reporting periods is greater when the interference variability is higher, i.e. when more users are active in the cell.

The Table 3 compares the performance of the three scheduling algorithms with a fixed number of users. The MaxCIR algorithm provides the best results in the pedestrian scenario. PF improves the RR performance since its functioning is channel state aware and hence it can take advantage of the good channel estimation. In spite of the greater fairness of PF as compared with MaxCIR, in this scenario with 30 users it is not enough to improve the MaxCIR performance in terms of delay. On the other hand, in the vehicular scenario all the algorithms suffer degradation in their performance, more pronounced in the case of the MaxCIR algorithm since this algorithm is totally channel state dependant and in the vehicular environment the channel estimation present lower accuracy. Again the degradation with increasing reporting periods is lower in the vehicular scenario for all the algorithms.

C. Performance with RT traffic.

To measure the level of QoS for the H.263 users, the user satisfaction concept (US) is introduced representing the percentage of H.263 frames transmitted before the next video frame is generated. The scenario for RT traffic considers a fixed number of 8 H.263 users transmitting at 64 kbps and moving in a cell with a radius of 2800m. Figure 1 shows the results after 2 hours of simulation for RR and MCIR algorithms. The set of simulated CQI reporting periods is 2, 8, 16, 32 and 64 ms.

From the results it can be concluded that the MaxCIR algorithm outperforms the functioning of the RR for the simulated environment in both the pedestrian and the vehicular scenarios. It is worth noting that the user satisfaction decreases as the reporting period increases but in a different way depending on the mobility. In the pedestrian scenario the slope is more pronounced than in the vehicular scenario.

TABLE II. NORMALIZED DELAY FOR MAXCIR ALGORITHM

N. of Users	CQI reporting period (ms)	pedestrian MCIR		vehicular MCIR	
		mean	95%	mean	95%
20	2	11.13	23.01	10.30	41.85
	16	13.56	26.83	11.10	44.16
	64	15.53	31.27	11.60	49.61
30	2	15.84	29.36	15.60	70.34
	16	17.63	34.22	15.47	71.15
	64	20.07	39.00	17.85	84.50
50	2	27.53	45.16	24.35	126.90
	16	28.86	48.97	28.10	144.40
	64	30.24	51.02	30.62	156.45

TABLE III. 95TH PERCENTILE OF THE NORMALIZED DELAY

		pedestrian			vehicular		
N. of Users	CQI reporting period (ms)	MCIR	PF	RR	MCIR	PF	RR
30	2	29.36	32.99	36.55	70.34	51.90	44.89
	16	34.22	34.52	42.43	71.15	52.03	45.92
	64	39.00	45.12	49.22	84.50	52.98	48.32

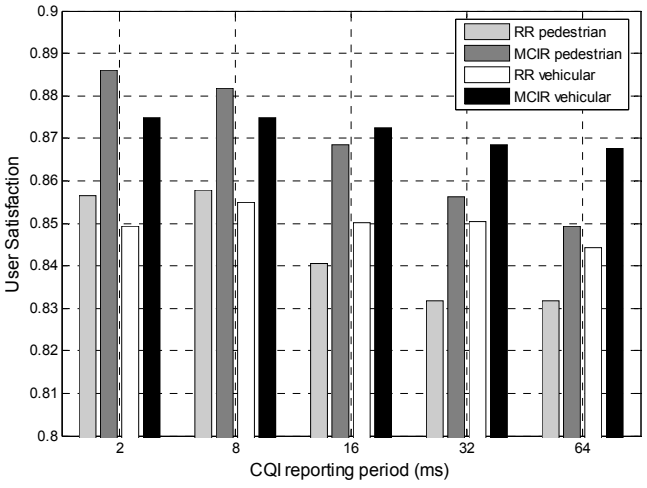

Figure 1. User satisfaction vs CQI reporting period with 8 H.263 users at 64 kbps for RR and MaxCIR algorithms

TABLE IV. USER SATISFACTION WITHOUT SERVICE DIFFERENTIATION

	MCIR		PF		RR	
CQI reporting period (ms)	H263	WWW	H263	WWW	H263	WWW
2	0.8247	0.9715	0.8155	0.9623	0.7847	0.9540
16	0.7965	0.9693	0.7789	0.9648	0.7521	0.9512
32	0.7885	0.9654	0.7705	0.9559	0.7367	0.9513
64	0.7762	0.9672	0.7296	0.9488	0.7325	0.9434

TABLE V. USER SATISFACTION WITH SERVICE DIFFERENTIATION

	MCIR		PF		RR	
CQI reporting period (ms)	H263	WWW	H263	WWW	H263	WWW
2	0.8741	0.9305	0.8624	0.9040	0.8464	0.8969
16	0.8749	0.9221	0.8614	0.9070	0.8423	0.8774
32	0.8570	0.9124	0.8604	0.9035	0.8320	0.8801
64	0.8636	0.9033	0.8579	0.8932	0.8211	0.8682

C. Performance in a mixed traffic scenario

Finally a mixed traffic scenario is considered with 8 H.263 users and 8 web users. Only pedestrian users are considered. The cell radius is again 2800 meters.

The user satisfaction in WWW is defined as the percentage of web pages transmitted in less than 4 seconds. Tables 4 and 5 collect the user satisfaction experimented by the H263 users and the web users in the simulated scenarios. From these results it can be concluded that the MaxCIR algorithm presents the best performance and RR shows again the worst functioning. The PF represents an intermediate point between these two algorithms since it allocates resources in a more intelligent way as compared with RR but does not take the best advantage from good channel conditions as MaxCIR does.

In the differentiation of services there is a transfer in the user satisfaction from the less priority service to the more priority service. By means of this differentiation it is possible to obtain better results in terms of H263 user satisfaction at the expense of a poorer WWW user satisfaction. The results show clearly that the service differentiation becomes a good option to maintain the H263 QoS. An improvement between 5 and 10 percentage points in the H.263 user satisfaction is observed. It is worth noting that the H263 user satisfaction is in this scenario worse than the observed in Fig. 1 since in this case some additional interference is produced in the system due to the transmission of web users.

VI. CONCLUSIONS

In this paper it has been evaluated the effect of the channel estimation inaccuracy on the performance of an HSDPA system. As explained, this inaccuracy is due to the delays in the acquisition and processing of the CQI reports. Generally if the reporting period of the CQI increases the system performance becomes worse. However, it has been demonstrated that this degradation depends on the speed of UE and in some situations on the load or on the resource allocating scheme. There is a trade off between the performance improvement produced by frequently reported CQIs and the degradation in uplink interference produced by these reports. Therefore, it is possible to optimise the CQI reporting period to reduce the number of CQI transmissions while maintaining the users QoS and the overall system performance.

ACKNOWLEDGMENT

This work has been partially funded by CICYT (Spanish National Science and Technology Council) and the FEDER program of the European Commission under the project TEC2005-08211-C02.

REFERENCES

[1] H.Holma and A.Toskala, HSDPA/HSUPA for UMTS. John Wiley and Sons, Ltd, 2006.

[2] F. Brouwer et alt., “Usage of Link-Level Performance Indicators for HSDPA Network-Level Simulations in E-UMTS”, in Proc. IEEE ISSSTA, Aug. 2004.

[3] Soo-Yong Jeon; Dong-Ho Cho, “Channel Adaptive CQI Reporting Schemes for HSDPA Systems”, IEEE Communications Letters, Vol.10, Issue 6, pp 459-461, Jun. 2006.

[4] Soo-Yong Jeon; Dong-Ho Cho, “An Enhanced Channel-Quality Indication (CQI) Reporting Scheme for HSDPA Systems”, IEEE Communications Letters, Vol. 9, Issue 5, pp 432-434, May. 2005.

[5] P.J.A. Gutiérrez, “Packet Scheduling and Quality of Service in HSDPA”, Ph.D.dissertation, Aalborg University, Oct. 2003.

[6] T. Kolding. Link and system performance aspects of proportional fair scheduling in WCDMA/HSDPA. IEEE Proc. VTC 2003 vol 3, pp 1717--1722, October 2003.

[7] M. Döttling, A. Seeger, M. Sikora, “Impact of Imperfect Channel Quality Feedback and User Terminal Capabilities on Throughput of HSDPA”. COST 273, Barcelona, Spain, January 2003.

[8] 3GPP, Physical layer procedures (FDD), 3GPP TS25.214 V7.1.0.

[9] J. Gozalvez y J. Dunlop, "System Performance and Adaptive Configuration of Link Adaptation Techniques in Packet-Switched Cellular Radio Networks", Computer Networks Journal (Elsevier), pp 404-426, 2006.

[10] W. Bang, K.I. Pedersen, T.E. Kolding, P.E. Mogensen, “Performance of VoIP on HSDPA”, IEEE Proc. VTC, Stockholm, June 2005.

[11] NEWCOM Project FP6-IST-507325-2002. Available at http://newcom.ismb.it.

[12] P. Barford and M. Crovella, “Generating representative web workloads for network and server performance evaluation”, Joint International Conference on Measurement and Modeling of Computer Systems, June 1998.

[13] O. Lázaro, D. Girma and J. Dunlop, “H.263 video traffic modelling for low bit rate wireless communications”, IEEE PIMRC, pp. 2124-2128, September 2004.

SPHERE – A Simulation Platform for Heterogeneous Wireless Systems

J. Gozálvez[1], D. Martín-Sacristán[2], M. Lucas-Estañ[1], J.F. Monserrat[2], J.J. Gonzalez-Delicado[1], D. Gozálvez[2] and M. Marhuenda[2]

[1]Signal Theory and Communications Division
University Miguel Hernández
Elche, Spain
j.gozalvez@umh.es

[2]iTEAM Research Institute, Mobile Communication Group
Polytechnic University of Valencia
Valencia, Spain
jomondel@dcom.upv.es

Abstract— **This paper presents SPHERE, a Simulation Platform for HEterogeneous wiREless systems, and describes its motivation, methodology and implementation approach. This advanced system level simulation platform emulates simultaneously the transmission of GPRS, EDGE Multi-slot, HSDPA and WLAN at the packet level, which allows conducting novel investigations on common radio resource management for beyond 3G systems or on the optimization of radio resource management techniques. This paper presents the simulation platform, validates it and introduces its research potential.**

Heterogeneous Access Networks, Simulation Platform, System Level, Common Radio Resource Management

I. INTRODUCTION

Once commercial 3G deployments are well on their way, research activities on the definition of beyond 3G (B3G) or fourth generation (4G) systems have started for a few years. There is a strong research consensus in that B3G or 4G systems will be characterized by the integration and joint management of various Radio Access Technologies (RATs), including current 2G/3G/3.5G cellular networks, WLAN, broadcasting systems and any potential new technology that might appear in the future. In this context, an important challenge in the path towards B3G heterogeneous wireless systems is to guarantee the interoperability and efficient management of the different RATs with the aim to guarantee the required Quality of Service (QoS) level and increase system capacity. As a result, strong efforts are being undertaken in the research community to define and optimize a Common Radio Resource Management (CRRM) framework [1].

The increasing complexity of current and future mobile wireless technologies requires the implementation of adequate platforms to evaluate and optimize their performance. Before considering a prototype or full-scale deployment, the use of simulation platforms is becoming increasingly common within the research community due to its cost/benefit ratio. However, it is important that to conduct meaningful and appropriate studies, such simulation platforms implement accurately the entities and process under evaluation. The implementation of such advanced simulation tools has become a very challenging task when investigating CRRM techniques since different RATs need to be simultaneously emulated in a single platform. Different research projects have been looking at the development of heterogeneous simulation platforms. For example, the work reported in [2] investigated different traffic distribution policies among a variety of RATs. However, the work was conducted at the session level and the simulation platform employed did not model each RAT's radio interface or the radio propagation effects. As a result, although the simulation platform allowed an investigation on load balancing schemes, it was not able to accurately determine final user perceived QoS values given its limitation on the radio modeling side. Related investigations on traffic distribution algorithms in heterogeneous wireless networks were presented in [3]. To conduct their investigations, the authors implemented a simulation tool that takes into account propagation models and some specific RAT-features although it does not model the complete radio transmission process. A different approach was considered in [4] where the authors considered an analytical model that relates the experienced CIR to the user perceived throughput. Nevertheless, the increasing complexity of mobile and wireless communication systems increases the difficulty of studying the performance of new techniques through analytical models. In fact, analytical studies usually require many simplifications and approximations that limit the accuracy and reliability of the obtained results. Important European projects, such as EVEREST, have also conducted advanced research work on heterogeneous wireless systems and CRRM techniques (EVEREST considered a GERAN-UMTS-WLAN scenario). This project has developed an interesting real-time emulator to demonstrate potential benefits of CRRM mechanisms [5]. This approach differs from the considered in this project where interest is place on system performance of CRRM techniques. Another significant heterogeneous wireless project is WHYNET project [6]. This project intends to develop a wireless hybrid network testbed to assess cross-layer interactions in heterogeneous wireless systems. However this project is more centered on sensor and mesh networks rather than focusing on CRRM in cellular systems.

The previous discussion has highlighted the availability of a variety of simulation tools devoted to conduct research on heterogeneous wireless systems. As explained, the modeling detail of these tools strongly depends on the type of work being conducted. Although each presented simulation platform is valid within their research framework, to the best of the authors' knowledge, there is not any simulation platform

available that implements at the packet/slot level different RATs and enables their simultaneous and parallel emulation.

In this context, the University Miguel Hernández and the Polytechnic University of Valencia have developed a novel, ambitious and scalable radio simulation platform for heterogeneous wireless systems, named SPHERE (Simulation Platform for HEterogeneous wiREless systems), under a common national research project. The platform currently integrates four advanced system level simulators, emulating the GPRS (General Packet Radio Service), EDGE (Enhanced Data-rates for GSM/Global Evolution), HSDPA (High Speed Downlink Packet Access) and WLAN RATs. The unique simulation platform emulates all four RATs in parallel and at the packet level, which enables an accurate evaluation of the final user perceived QoS through the implementation of novel CRRM and RRM mechanisms. The radio interface specifications of these four technologies have been faithfully implemented in the SPHERE simulation platform, which works with a high time resolution (in the order of some milliseconds). This modeling approach validates the capability of the SPHERE simulation platform to dynamically and precisely evaluate the performance of RRM/CRRM techniques. The platform has been developed following a modular and scalable design, which guarantees an easy adaptation of the platform configuration to specific requirements, and allows the rapid integration of new RATs. In particular, the research team is considering the future expansion of the platform to also emulate the UMTS and Mobile WiMAX radio interfaces.

The interest of the SPHERE platform and the research being conducted is highlighted by the support of important companies in the mobile and wireless industries, such as Motorola, Swisscom Innovations and Telefonica I+D, to the research project developing SPHERE.

II. RADIO ACCESS TECHNOLOGIES

As it has been previously said, the SPHERE simulation platform currently emulates GPRS, EDGE, HSDPA and WLAN transmissions. This section briefly summarizes the main characteristics of these radio interfaces with regard to SPHERE.

A. GPRS

The GPRS radio interface is based on a combined FDMA/TDMA multiple access mechanism and a FDD scheme. The GPRS standard can be modeled as a hierarchy of logical layers with specific functions. Prior to transmission, data packets are segmented into smaller data blocks across the different layers, with the final logical unit being the Radio Link Control block which has a duration of 20ms. The resulting RLC data blocks are then coded and block-interleaved over four normal bursts in consecutive TDMA frames. Although GPRS is based on a single modulation scheme it defines four different coding schemes (see Table I.) that have all been emulated within SPHERE.

A GPRS TDMA frame is equal to 4.615 ms and is divided into eight 0.577 ms time-slots. Such time-slots impose the SPHERE time resolution for the GPRS radio interface. GPRS defines a temporal hierarchy with higher order structures such as super- and hyper-frames that have not been implemented in SPHERE since the platform is aimed at radio resource management investigations.

TABLE I. GPRS TRANSMISSION MODES

Mode	Modulation	Code Rate	Bits per Radio Block	Bit Rate (kbps)
CS-1	GMSK	1/2	181	9.05
CS-2	GMSK	≈ 2/3	268	13.4
CS-3	GMSK	≈ 3/4	312	15.6
CS-4	GMSK	1	428	21.4

B. EDGE

The EDGE radio interface is based on the same multiple access scheme as GPRS, but considers different transmission modes (Modulation and Coding Schemes, MCS), all implemented in the SPHERE platform following the description in Table II. The main difference to GPRS is the introduction of 8PSK, a multilevel modulation that theoretically increases EDGE data rates by a factor of three.

TABLE II. EDGE TRANSMISSION MODES

Mode	Modulation	Code Rate	Family	Bits per Radio Block	Bit Rate (kbps)
MCS-1	GMSK	0.53	C	1×176	8.8
MCS-2	GMSK	0.66	B	1×224	11.2
MCS-3	GMSK	0.85	A pad.	1×272	13.6
			A	1×296	14.8
MCS-4	GMSK	1.00	C	2×176	17.6
MCS-5	8-PSK	0.37	B	2×224	22.4
MCS-6	8-PSK	0.49	A pad.	2×272	27.2
			A	2×296	29.6
MCS-7	8-PSK	0.76	B	4×224	44.8
MCS-8	8-PSK	0.92	A pad.	4×272	54.4
MCS-9	8-PSK	1.00	A	4×296	59.2

The EDGE transmission modes are divided into three different families, namely A, B and C. Each family has a different basic payload unit of 37 (and 34), 28 and 22 octets respectively. Different code rates within a family are achieved by transmitting a different number of payload units within one radio block. For families A and B, 1, 2 or 4 payload units can be transmitted per radio block, while for family C, only 1 or 2 payload units can be transmitted. These families are designed to allow a radio block to be retransmitted with a transmission mode, within the same family, different from that used in the original transmission; this option is not possible in the current GPRS standard. A block received in error can be resegmented and retransmitted using a more robust transmission mode within the same transmission family.

The GPRS and EDGE transmission procedures are very similar, although some differences for high order modes are appreciated. When 4 payload units are transmitted (MCS-7, MCS-8 and MCS-9), these are split into two separate blocks. These blocks are in turn interleaved over only two bursts, for MCS-8 and MCS-9, and over four bursts for MCS-7. All the other MCSs can only transmit a single block that is interleaved over four bursts. When switching to MCS-3 or MCS-6 from MCS-8, three or six padding octets are, respectively, added to

fill a radio block. Identically to GPRS, the transmission of a whole EDGE radio block requires 20 ms.

C. HSDPA

HSDPA is based on a CDMA multiple access scheme and considers both a FDD and TDD component, although SPHERE only emulates the FDD one. The FDD mode operates at a chip rate of 3.84 Mcps, which results in an approximated bandwidth of 5 MHz. In the time domain, a Transmission Time Interval (TTI) of 2 ms is defined. A TTI is further divided into three 667 μs slots. In the code domain, channelization codes at a fixed spreading factor of 16 are used. Multi-code transmission to a single user during a TTI is also allowed.

HSDPA achieves high data rates of up to 14 Mbps by means of adaptive modulation and coding (AMC), fast scheduling mechanisms (each TTI) and a powerful Hybrid ARQ mechanism. AMC or Link adaptation (LA) is a process of paramount importance to optimize system functioning. Its operation is based on user equipment reporting the channel state either cyclically or in a triggered-based manner by means of the Channel Quality Indicator (CQI). The definition and processing of the CQIs is explained in detail in [7]. Numerically, CQI varies from 1 to 30, increasing its value when the channel quality augments. To model the radio channel quality, the simulations reported in this paper considered several look-up tables (LUT) matching the SINR as a function of the BLER; in particular, one for each CQI such that the maximum CQI can be calculated considering a specific QoS. These LUT also include the effect of the HARQ retransmission with chase combining.

The available number of codes has also been carefully taken into account and, in the same way, power consumption of all the control channels has been considered to determine the available power per user. Assuming code multiplexing of n users per TTI, n HS-SCCH channelization codes should be allocated, whereas the available power is equally divided among the n users. The maximum number of HS-SSCH codes has been set to 4.

D. WLAN

Current WLAN standards do not contemplate the same level of radio resource management functionality than mobile systems. However, extensions that support a more advanced RRM framework have been developed in standardization bodies. In this context, SPHERE implements the 802.11e specification, which provides more advanced MAC mechanisms to support QoS. This standard specifies two access mechanisms, the Enhanced Distributed Channel Access (EDCA) and the HCF controlled channel access (HCCA). According to the literature, the optimum system operation corresponds to the case in which both access mechanisms work together, and this is the philosophy followed in SPHERE.

At physical layer both WLANs 802.11b/g versions have been implemented. Table III summarizes the list of properties for both specifications. In SPHERE, user equipments are simply characterized by the receiver sensitivity (S) and the transmission power which has been set to 100 mW (20 dBm).

TABLE III. WLAN PHY MAIN CHARACTERISTICS

Scheme	Mod.	Max PHY Data Rate (Mbps)	Max MAC Data Rate (Mbps)	Family	S (dBm)
DSSS	BPSK	1	0.8	.11b	-94
	QPSK	2	1.2	.11b	-92
	CCK	5.5	3	.11b	-91
	CCK	11	5.4	.11b	-89
OFDM	BPSK	6	4.1	.11g	-82
	BPSK	9	5.8	.11g	-81
	QPSK	12	7.1	.11g	-79
	QPSK	18	9.4	.11g	-77
	16QAM	24	11.0	.11g	-74
	16QAM	36	13.3	.11g	-70
	64QAM	48	16.8	.11g	-66
	64QAM	54	17.8	.11g	-65

III. THE SPHERE PLATFORM

Fig. 1 shows the scenario modeled by the SPHERE platform which includes the GPRS, EDGE, HSDPA and WLAN radio interfaces. As shown in Fig. 1, the SPHERE platform does not only model the radio interface of the four technologies but also implement various RAT specific RRM features and a centralized CRRM entity. This entity directly collects specific RAT information (e.g. load, channel quality conditions, etc) and interacts with the RRM entities implemented at each RAT.

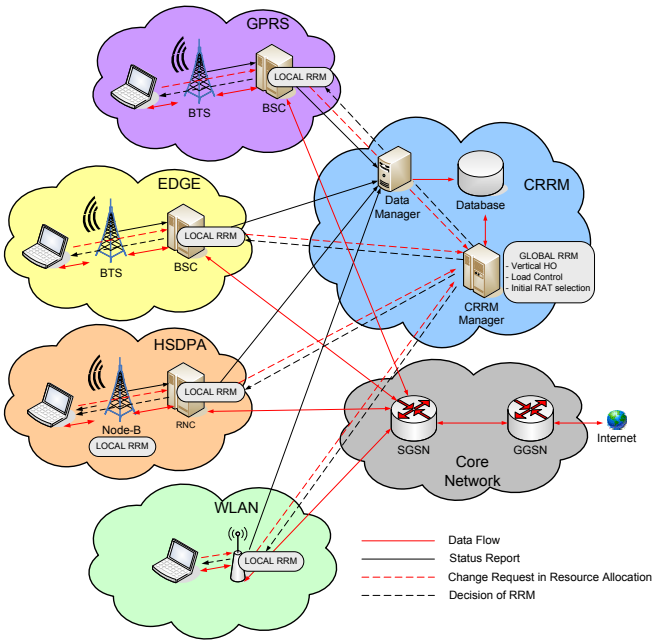

Figure 1. SPHERE heterogeneous scenario

A logical structure of the SPHERE simulation platform, which is a discrete-event system level simulator concentrated on the downlink performance, is shown in Fig. 2. The components shown in this figure, their features, interactions and data flow will be described in the following sections. Finally, the potential of the platform to conduct advanced research on the design, evaluation and optimization of CRRM and RRM techniques will be demonstrated.

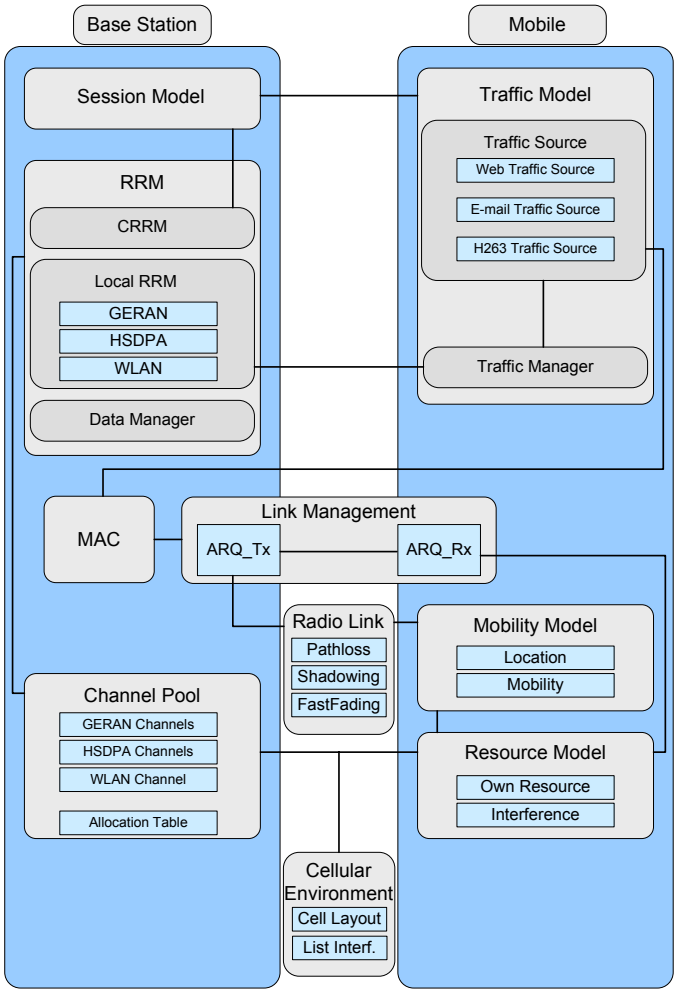

Figure 2. SPHERE logical structure

A. *Cellular Environment*

The *Cellular Environment* entity is a system module storing the location of each base station and the interfering relations among them; this information is needed to estimate the experienced interference levels. The cellular layout can be modified offline at any time to change the system configuration under study. Currently, the SPHERE platform considers a cell layout of 27 omni-directional cells. In order to avoid border effects, a wrap-around technique has been applied. Three-sectorized cell sites have also been modeled. A basic scenario has been defined considering two concentric coverage cells with increasing cell radii. The smaller one, of 50m, represents the coverage area of the WLAN technology. The second coverage radius of 500m includes HSDPA, EDGE and GPRS.

B. *Radio Link*

This module models the radio propagation conditions between transmitter and receiver and is a generic entity employed by any wireless link established. It characterizes the three radio propagation effects, namely path loss, shadowing and fast fading.

The path loss model provides an average measure of the signal attenuation over a given distance. For cellular systems, the path loss (in dB) reported in [8] has been considered:

$$L_p = \left(40 - 16 \cdot 10^{-2} \Delta h_b\right) \log_{10} d - 18 \log_{10} \Delta h_b + 21 \log_{10} f + 80 \quad (1)$$

Where d is the distance in km between the base station transceiver and the mobile terminal, f is the carrier frequency in MHz and Δh_b is the base station antenna height, measured in meters from the average roof top level. In HSDPA typical values for these parameters are Δh_b=15m and f=2000 MHz, leading to

$$L_p\left(dB\right) = 128.1 + 37.6 \log_{10} d \quad (2)$$

This model is applicable to frequencies ranging from 1.5 GHz up to 2 GHz. As a result, a carrier frequency of 1.8 GHz has been assumed for GPRS and EDGE. This change in the frequency parameter results in the following path loss equation

$$L_p\left(dB\right) = 127.2 + 37.6 \log_{10} d \quad (3)$$

For WLAN, the following path loss model has been implemented [9]:

$$L_p(dB) = 145 + 35 \log_{10} d \quad (4)$$

Various works have demonstrated the importance of an adequate shadow modeling to conduct appropriate RRM investigations [10]. In fact, the shadowing effect results in additional signal attenuation due to obstacles in the path between transmitter and receiver. Measurements have shown that the shadowing loss can be modeled as a random process with a normal distribution of mean 0 dB and standard deviation between 4 and 12 dB depending on the propagation environment. The SPHERE platform considers an urban or suburban environment and, as a result, a shadowing standard deviation equal to 6 has been set. The shadowing is a spatially correlated process so that the shadowing loss experienced by a mobile at a given position is correlated to that experienced at a nearby position. Although the authors are actually working on migrating the shadowing models to that used in [10], the current version of the SPHERE platform models this spatial correlation as detailed in [11], with a de-correlation distance of 20 m.

Fast fading modeling is also important when considering technologies, such as EDGE and HSDPA, that base their radio operation on link adaptation techniques [12]. In these RATs, transmission conditions are modified depending on the current channel characteristics. For the sake of simplicity, in SPHERE only a simple block fading model is considered, i.e. the fast fading stays constant over a coherence time interval and each sample is statistically independent. Hence, in addition to the path loss and shadowing value of each radio block, in SPHERE a third multiplicative factor is considered when determining the received carrier: the fast fading coefficient. This factor is of unit mean and follows the probability density function:

$$f_\psi\left(\psi\right) = \frac{M}{\left(M-1\right)!} \left(M\psi\right)^{M-1} e^{-M\psi} \quad (5)$$

Here, M denotes the number of resolvable independent multi-paths at the receiver. In GPRS and EDGE M is set to 1 and in HSDPA M=3.

C. Base Station

As shown in Fig. 2, the Base Station entity is responsible for the Medium Access Control (MAC) and RRM functions. It also controls the channel pool where the status of all channels per RAT is maintained. In the Base Station is also located the session generation process. Once a new mobile is active in the system, the CRRM entity chooses its initial RAT depending on a specific policy. When a mobile station requests a channel from a given RAT, the channel pool of the serving base station is examined to search for an available channel. If a free channel is available on the requested RAT, the mobile station is assigned a randomly chosen channel or based on some quality metrics [13]. If a free channel is not available on the requested RAT, the mobile station is assigned a channel from a different RAT, depending on the CRRM scheme under consideration, or it is placed in a queue until a transmitting mobile ends its transmission and releases its channel. For users in GPRS and EDGE queues, a First-Come First-Served (FCFS) scheduling policy is applied so that channel requests are satisfied in the same order as they appear. Users in the HSDPA queue can be served either in a round robin fashion, according to the Max C/I criterion, which selects at any moment the user with better transmission quality, or following the proportional fair algorithm. In WLAN, real-time traffic is delivered through HCCA with a FCFS policy, whereas best effort users mutually contend to get the channel control being served using the EDCA protocol.

Apart from the scheduling, other implemented RRM functionalities include Link Adaptation for GPRS, EDGE, HSPDA and WLAN, multi-channel operation for GPRS and EDGE, multi-code allocation in HSDPA and call admission control in all technologies.

Link adaptation (LA), also referred to as Adaptive Modulation and Coding (AMC) in HSDPA, is an adaptive RRM technique that periodically estimates the channel quality conditions and selects the optimum transmission mode based on a predefined selection criterion. For web browsing and email services, the transmission mode that maximizes the throughput is selected. For H.263 video service, in GPRS and EDGE the algorithm proposed in [14] has been used since it outperforms the former in several key aspects affecting real-time operation. In the case of multi-channel transmissions, the channel quality conditions are estimated over all channels simultaneously assigned to a single user and their average value is used to estimate the optimum transmission mode according to the established selection criterion. In HSDPA, the mobile directly reports its channel conditions to the base station by means of the CQI. With this information, the base station knows the maximum allocable number of codes as well as the modulation and coding scheme. The final allocation shall always be as higher as possible but not exceeding the transmission mode reported by the user. Automatic Rate Fallback (ARF) is the implemented LA algorithm for WLAN; ARF and other algorithms with similar operating concepts have been widely implemented in many WLAN products although they are not included in the IEEE standards. In ARF, the sender deduces the channel conditions by measuring the numbers of consecutively successful and failed transmissions. The sender adjusts its modulation mode and data rate in accordance with these measurements.

Multi-channel or multi-code mechanisms based on some innovative schemes currently under investigation have been incorporated into the SPHERE simulation platform for the GPRS-EDGE and HSDPA radio interfaces. In these schemes, the number of channels or codes that a base station can simultaneously allocate to a single user depends on factors such as the capability of the terminal, the system load, the availability of radio resource, the requested service type and the considered multi-channel allocation policy.

The research team behind the SPHERE simulation platform is currently working on the development of CRRM schemes. Such schemes, briefly analyzed in section VI, base their RAT selection on utility functions and operating parameters, such as the RAT load, required service and its QoS parameters, interference levels and effect on active transmission, etc. The authors are also working on adaptive CRRM schemes interacting with a RAT RRM functions to compensate or penalize inadequate CRRM decisions. In the current implementation, a RAT selection can be performed for each new session, periodically, or every time a new packet is generated. It is important to highlight that RAT changes are done dynamically in the SPHERE platform and that the radio transmission can be immediately resumed with the newly selected RAT at the stage where the radio transmission ended using the previous RAT. The platform has also been prepared to consider the case in which a user handles different application sessions through various RATs.

D. User traffic behavior

User traffic demands are usually described at two levels: session-arrival process and traffic models. Session-arrival processes, also referred as traffic generation, are usually modeled as a birth-death process, which can be characterized by the following parameters: busy hour call attempts (BHCA), arrival distribution, mean session duration, duration distribution, etc. On the other hand traffic models describe the source behavior within a session. They vary depending on the type of service and they can be described by parameters such as: average active/inactive times, time distributions, data rate, packet length distribution, etc. In SPHERE, the session-arrival has been implemented at the base station, while the traffic models are controlled by the mobile station for optimizing the code.

1) Session Model

Three different services have been implemented in SPHERE, namely web browsing, real-time H.263 video transmissions and email. Cellular subscribers are usually considered to have independent behavior one from each other, which results in exponentially distributed inter-session arrival times. For each one of the implemented services a specific inter-session time is defined, which allows controlling the traffic load.

2) Traffic Model

Despite considering downlink transmissions, the current version of the SPHERE platform implements the traffic models at the mobile station. This has been done to optimize the simulation code.

The web browsing service follows the model described in [15]. It follows an ON/OFF pattern where a web browsing session starts with the submission of a web page request by the user. The time interval needed to transfer the requested web page is referred to as active period. When the transfer is completed, the user will take some time to read the information before initiating another request. This time corresponds to the inactive period. The implemented model is based on the HTTP 1.0 standard where a different TCP connection is established for the transmission of each object in a web page. In this case, the active ON time has been considered as the time needed for the transmission of a single object of a web page, while the active OFF time represents the time elapsed between closing a TCP connection and opening a new one to transfer another object of the same web page. The implemented web browsing traffic distributions are shown below:

TABLE IV. PARAMETERS OF THE WEB BROWSING TRAFFIC MODEL

Parameter	Mathematic Distribution	Probability Distribution Function	Constants
Object Size	Pareto	$f(x)=\frac{\alpha k^{\alpha}}{x^{\alpha+1}}$	k=1000 α=1,0
Active OFF Time	Weibull	$f(x)=\frac{b\,x^{b-1}}{a^{b}}e^{-(x/a)^{b}}$	a=1,46 b=0,382
Inactive OFF Time	Pareto	$f(x)=\frac{\alpha k^{\alpha}}{x^{\alpha+1}}$	k=1 α=1,5
Number of Objects per Web Page	Pareto	$f(x)=\frac{\alpha k^{\alpha}}{x^{\alpha+1}}$	k=1 α=2,43

The email traffic model also follows an ON/OFF pattern [16]. The model assumes that incoming messages of a user are stored at a dedicated email server. This server keeps the emails in a mailbox until the user logs onto the network, following the session model previously described, and downloads the emails. When the user opens the mailbox, the headers of the available messages are downloaded. The user scans then through these headers and downloads the emails she/he is interested in. When the user finishes downloading a message (active period), she/he will read it (inactive period) before downloading the next message, and so on. The employed email traffic distributions are shown below.

Real-time services have also been included in SPHERE through the emulation of real-time H.263 video transmissions following the model presented in [17]. This model takes into account the three different frame types considered in the H.263 standard, namely I, P and PB. The model characterizes the size and duration of the video frames, the correlation between both parameters for each video frame, and the transition probability between different video frame types. The modeling is performed at two levels. The first one establishes the frame type to generate. I-frames are periodically created, while a Markov chain drives the transition generation between P- and PB-frames. Once the frame type is selected, the model determines the size and the duration of the video frame to be transmitted. The reader is referred to [17] for a detailed analytical described of the real-time H.263 video traffic model.

TABLE V. PARAMETERS OF THE EMAIL MODEL

Parameter	Mathematic Distribution	Probability Distribution Function	Constants
Email Size	Weibull	$F(x)=\begin{cases}1-e^{-e^{-k_1 x c_1}}, & F(x)\le 0,5\\ 1-e^{-e^{-k_2 x c_2}}, & F(x)>0,5\end{cases}$	k_1=17,6 c_1=3,61 k_2=2,04 c_2=0,37
Inactive Period	Pareto	$F(x)=1-\left(\frac{k}{x}\right)^{\alpha}$	k=30 α=0,5

E. Mobile station

1) Mobility Model

The implemented mobility model considers a suburban scenario where users move at constant speed. The initial position of a mobile station within a cell is randomly set according to a random uniform distribution. The discrete nature of event-driven simulations has been reflected in the mobility model through the definition of the time at which a mobile's movement is updated. The length of each step is constant and equal to the decorrelation distance used for the shadowing model; the position of a mobile at a particular time between two random positions is extracted by interpolation. The direction of each step is randomly established by adding a random angle to the previous direction. The random angle is obtained from a normal distribution with zero mean and a variance dependent on the mobile speed. The mobility model, leading to a long-term uniform user's density within a cell, was shown to be consistent with an analysis performed on real data provided by a mobile operator.

2) Resource Model

The resource model entity is basically responsible for controlling the radio transmission parameters of a channel currently assigned to a user and for estimating the experienced channel quality conditions, in this case carrier to interference ratio. For WLAN systems only the received power is calculated since it is the only value needed to obtain the maximum allocable bit rate.

The GPRS and EDGE CIR level is estimated as follows:

$$CIR_{GPRS/EDGE}=\frac{\frac{P_i}{L_P^{ii}\cdot L_S^{ii}}\psi_i}{\sum_{j\in\Omega}\frac{P_j}{L_P^{ij}\cdot L_S^{ij}}\psi_j+N_0\cdot W} \qquad (6)$$

where P_i is the power transmitted by the reference cell (cell i) to the user of concern, L_P^{ii} and L_S^{ii} are the path loss and shadowing loss over the link between transmitter and receiver at the reference cell, Ω is the set of active co-channel interferers, P_j is the transmission power of each one of the interfering channels, L_P^{ij} and L_S^{ij} are the path loss and

shadowing loss over the link between the active transmitting interferers in cells *j* and the interfered receiver at the reference cell *i*, and $N_0 \bullet W$ represents the thermal noise at the receiver in the reference cell, with N_0 being the noise spectral density and W the bandwidth of the transmission channel. Finally ψ models the fast fading effect.

In CDMA-based systems, such as HSDPA, channelization codes for the users of the same cell are perfectly orthogonal. However, due to multi-path fading, this orthogonality decreases and some intra-cell interference component is observed. Intra-cell interference on a CDMA system is modeled by an orthogonality factor [18], which is usually denoted as α. In absence of multi-path fading, the codes are perfectly orthogonal, so $\alpha = 1$. In the worst case $\alpha = 0$, meaning that orthogonality is entirely destroyed. Typical values of α are between 0.4 and 0.9. Thus, the HSDPA CIR level can be expressed as follows:

$$CIR_{HSDPA} = \frac{\frac{P_i}{L_P^{ii} \cdot L_S^{ii}} \psi_i}{\frac{(P_{Ti} - P_i/C_i) \cdot (1-\alpha)}{L_P^{ii} \cdot L_S^{ii}} \psi_i + \sum_{j \in \Omega} \frac{P_{Tj} \psi_j}{L_P^{ij} \cdot L_S^{ij}} + N_0 \cdot W} \quad (7)$$

where P_i is the addition of the power transmitted in all the C_i channels allocated by the reference cell to the user of interest, P_{Ti} is the total power transmitted by the reference cell, Ω is the set of cells interfering the user and P_{Tj} is the total power transmitted by these interferers. In this expression, the parameters P_{Ti} and P_{Tj} also include the base station power reserved for other channels different from the HSDPA High Speed Downlink Shared Channel (HS-DSCH). The power allocated to each control channel has been extracted from [19]. ψ also models the fast fading effect.

In WLAN the received power is calculated as:

$$P_{Rx} = \frac{P_i}{L_P^{ii} \cdot L_S^{ii}} \quad (8)$$

It is worth noting that the WLAN interference from the rest of Access Points has not been considered in the SPHERE platform since the distance between WLAN transmitters is high enough to neglect it.

F. Link Management

The Link Management module is responsible for handling the radio transmission and emulating channel errors.

1) Transmission Process

GPRS controls the radio transmission of RLC blocks through an Automatic Repeat reQuest (ARQ) protocol, described in specification 3GPP TS 04.60, that is implemented in SPHERE. This ARQ protocol is based on the numbering of the blocks and a selective repeat principle with sliding transmitting and receiving windows. The transmitter sends blocks and the receiver sends acknowledgment messages acknowledging correctly received blocks and requesting the retransmission of erroneously received blocks. The transmitting and receiving ARQ windows have a size of 64 RLC blocks. The reporting period, which defines how regularly the receiver sends acknowledgment messages, has been set to 16 blocks. No block losses and errors on the transmission of the acknowledgement messages have yet been considered in SPHERE. A similar ARQ protocol has been implemented for EDGE, with varying window sizes according to the number of channels simultaneously assigned to a single user; the window sizes range vary from 64 to 1024 radio blocks. For EDGE transmissions, a 32 radio blocks has been selected.

In HSDPA, retransmission of erroneous transport blocks is performed by an N-channel stop-and-wait (SAW) ARQ protocol. In stop-and-wait schemes, the transmitter handles the transmission of a single block until it has been successfully received. In SPHERE, a maximum of 8 channels can be set up simultaneously since this is the value suggested in the standard. Block size is determined by the reported CQI. As for GPRS and EDGE, no transport block losses or errors on the acknowledgement messages have been emulated.

For WLAN, the SPHERE platform also implements an ARQ protocol. In this case only one channel SAW is employed. The transport block is a fixed-length IP packet of 1500 bytes although it is possible to perform fragmentation to improve channel utilization.

2) Channel Errors Emulation

For GPRS and EDGE, in order to decide whether a radio block is received in error, the experienced CIR is computed in the four TDMA frames used to transmit such block. After completing the transmission of a whole radio block, the four associated CIR values are averaged and a single CIR_{avg} value is obtained. This CIRavg value is then mapped to a BLock Error Rate (BLER) value ($BLER_0$) by means of a Look-Up Table (LUT) such as the ones proposed in [12]. LUTs are used as a means of interfacing link and system level simulations using the link level analysis as a source of information for the system level. The link level performance is then represented by a simplified model consisting of a set of LUTs mapping the CIR to a given link quality parameter such as the BLER. Different LUTs need then to be produced for different operating conditions, e.g., transmission mode, mobile speed and propagation environments (typical urban or rural area). Once a $BLER_0$ value is obtained from the adequate LUT, a random process is used to decide whether the radio block is correctly received.

For HSDPA, the experienced CIR is computed in each slot of a TTI. After completing the transmission of a whole transport block, the three associated CIR values are averaged and a single CIR_{avg} value is obtained, which represents the quality experienced by the transport block. In the same way as in the GPRS-EDGE case, a LUT is employed to map the CIR value in a BLER value and to decide whether a block is correctly received. When a transport block is received in error, it is not discarded but stored in the receiver buffer and combined with retransmissions according to a specific method. When Incremental Redundancy (IR) is employed, retransmissions are typically not identical to the original transmission. Another possibility is to use the Chase Combining (CC) scheme in which retransmitted blocks are identical to that of the first transmission. After several transmissions, the resulting effective CIR value (CIR_{eff}) is

representative of the global quality experienced by the data stored in the receiver buffer after the combination process. The value of CIR_{eff} is used to decide whether the information is received in error. To this end, a random number X_0 between zero and one is drawn from a uniform distribution prior to the first transmission of a transport block. This random number is mapped to a CIR value by means of the corresponding LUT. The CIR value is then established as the minimum effective CIR (CIR_{min}) that must be obtained at the receiving side to consider that the information is correctly decoded. Each time a transport block is transmitted, the decision is taken by comparing CIR_{min} with the CIR_{eff} obtained after the combination of the current transmission and previous transmissions (if any) of the transport block. When $CIR_{eff} \geq CIR_{min}$, the transport block is assumed to be successfully received. However, if $CIR_{eff} < CIR_{min}$, the transport block is then assumed to be received in error.

When modeling WLAN it is unusual to make use of LUTs as in cellular systems. Rather, the concept of sensibility is employed. If a certain physical transmission mode is given and the mean received power is over its specific sensibility then the transmitted block will be properly received, otherwise the block is dropped.

IV. RRM AND CRRM INVESTIGATIONS

Once the SPHERE simulation platform has been described, this section is devoted to validate the implemented platform and show its potential for conducting novel CRRM and RRM investigations. In particular, this section briefly presents EDGE, HSDPA and WLAN RRM work conducted using the SPHERE platform and initial CRRM investigations considering a heterogeneous scenario made up of the GPRS, EDGE and HSDPA radio interfaces.

A. Platform Validation

The objective of this section is to validate the SPHERE platform by means of simulation results. For a rigorous validation, it would be desirable to compare the results obtained by SPHERE with the results from another source where a similar heterogeneous scenario was simulated. However, as no source for comparison has been found, this section compares the SPHERE results with the maximum possible performance of each RAT to prove the obtained results are within the expected range.

Figure 3 shows the Cumulative Distribution Function (CDF) of the throughput for the different RATs implemented in SPHERE. The EDGE RAT has been simulated considering different multi-slot configurations. These values have been obtained by simulating the diverse RATs of the system independently, i.e. not simultaneously, and considering a load of 15 users per cell in each simulation (3 for web-browsing, 3 for email, 3 for H.263 at 32 kbps, 3 for H.263 at 64 kbps, and 3 for H.263 at 256 kbps). In the simulation of WLAN, user's mobility is restricted to their 50m cell radius.

As it can be observed from Figure 3, the performance per RAT does not overpass their maximum theoretical value, which validates the current implementation of the SPHERE platform. In fact, the performance is lower than the maximum theoretical values given the low cell radius, high transmitting power and high cell load used to conduct these simulations. These conditions increase the interference levels which in turn increases the experienced BLER, decreases the throughput performance and promotes, in adaptive radio interfaces such as those modeled in SPHERE, the use of transmission modes with high error protection and lower bit rate. Figure 3 also shows that the HSDPA throughput performance is improved when using IR instead of CC. This improvement is due to the higher IR probability of successfully decoding retransmitted blocks given that it sends additional redundancy information with each retransmission.

Results clearly show that the SPHERE platform offers a considerably wide range of transmission capabilities, highlighting its suitability for analyzing CRRM policies in a heterogeneous wireless framework. From such perspective, and given that EDGE under multi-slot operation increases the transmission bit rate 'variety', each EDGE multi-slot configuration could be considered as a different RAT for the CRRM policies implemented in heterogeneous systems. This results in a heterogeneous wireless framework with a larger set of radio access alternatives that offer a wider set of transmission capabilities.

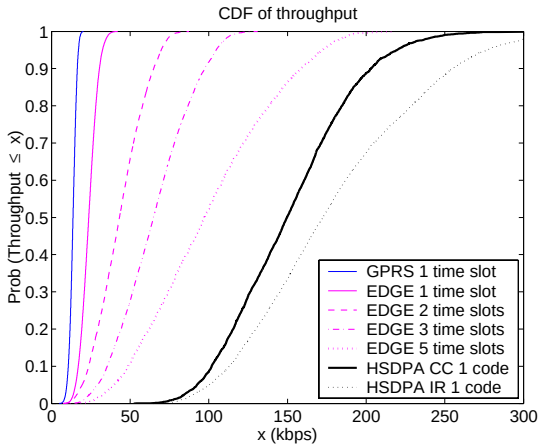

Figure 3. SPHERE throughput cdf

When analyzing the mean throughput performance, the results do not vary significantly per traffic service in a given RAT. Despite similar performance, user satisfaction is not similarly maintained for each RAT since each service has different QoS expectations. Consequently user satisfaction parameters per traffic service have been defined. Web-browsing and email users are assumed to be satisfied when they download a web page or an email in less than 4 seconds, as specified in 3GPP TS 22.105. Video users are supposed to be satisfied every time a video frame is entirely received before a new one is generated, i.e. no part of the video frame is discarded. The user satisfaction is therefore defined as the percentage of times that a web page, email o video frame transmission results satisfactory for the end user. Table VI shows the obtained results. In general, for a given service, the user satisfaction increases as the selected RAT offers better capabilities. However, it can be observed that for some H.263 users, EDGE using 5 slots obtains better user satisfaction than

HSDPA despite its lower throughput performance. This is due to two main reasons. First of all, it is important to note that the defined user satisfaction parameter for video transmissions is based on the bit rate and not on the throughput since transmission errors have not been accounted in the parameter's definition. Also given the real-time nature of H.263 video transmissions, retransmissions of erroneously received data blocks for EDGE and GPRS have not been allowed. On the other hand, up to four retransmissions have been allowed for HSDPA to take profit of HARQ capabilities. As a result, given that the experienced BLER is quite high for the considered operating conditions, HSDPA requests several retransmissions of a transport block and therefore less video frames are transmitted before the next one is generated than considering EDGE with 5 slots.

TABLE VI. USER SATISFACTION (%)

	Web	**Email**	**H.263 32 kbps**	**H.263 64 kbps**	**H.263 256 kbps**
GPRS	0.0	44.5	0.4	0.3	0.0
EDGE (1S)	53.8	53.9	87.9	43.0	0.0
EDGE (2S)	64.1	55.7	99.9	85.9	0.2
EDGE (3S)	74.4	59.7	100.0	98.9	0.6
EDGE (5S)	84.7	71.9	100.0	100.0	15.9
HSDPA	99.8	98.8	99.9	99.7	95.9
WLAN	100.0	100.0	100.0	100.0	100.0

As QoS requirements increase, RATs with higher bit-rates are required in order to obtain an acceptable degree of satisfaction. For the most demanding service, i.e. H.263 video transmission with a mean bit-rate of 256 kbps, only HSDPA and WLAN are able to offer an acceptable satisfaction level to the users for the considered operating conditions. On the other hand, for services with low QoS requirements such as background services, RATs with limited capabilities can fulfill the user expectations.

B. EDGE multi-channel operation

The authors are currently working on novel multi-channel operation mechanisms designed to effectively distribute the scarce available resources among users, according to the system state and the user's needs based on some contracted QoS and the service type requested. In particular, the authors are working on bankruptcy theory and utility-based schemes in a multi-service scenario. The techniques under evaluation include Discrete Constraint Equal Award (DCEA), Discrete Constraint Equal Loss (DCEL), maximum utility (MaxUtil) scheme and Required Data Rate (ReqDataRate).

In DCEA, resources are allocated first to the users experiencing lower satisfaction levels until all users reach the same user satisfaction (the same user satisfaction for different services might not require the same number of channels). In case an equal user satisfaction cannot be reached, priorized users (e.g. real-time H.263) receive the remaining channels. DCEL follows a similar assignment except that users get initially all necessary resources to guarantee maximum user satisfaction, and resources are sequentially reduced to try to guarantee at the end of the channel distribution the same user satisfaction levels. As for DCEA, if such aim is not possible, real-time services are priorized. In the MaxUtil scheme, channels are assigned to the users that will see a higher increase in user satisfaction if a channel is being assigned to them. The ReqDataRate mechanism calculates the required data rate to transmit pending data in time to maintain user satisfaction and allocates the channels necessary to reach such data rate, again based on service type priority.

An example of the performance of the four multi-channel allocation schemes currently being assessed using the SPHERE simulation platform is shown in Table VII.

TABLE VII. USER SATISFACTION (%) FOR MULTI-CHANNEL ASSIGNMENT SCHEMES IN EDGE

Service	**DCEA**	**DCEL**	**MaxUtil**	**ReqDataRate**
Web	48.94	57.48	45.52	46.38
Email	11.31	15.09	91.60	10.09
16 kbps	86.52	63.71	99.93	89.91
32 kbps	87.54	79.64	82.58	88.54
64 kbps	87.18	83.65	86.34	85.15

C. Effect of Channel-Quality Indicator Delay on HSDPA

The authors have also assessed the intrinsic delay of the Channel Quality Indicator (CQI) reporting process and its effect on the HSDPA system performance. As the authors are currently investigating, this delay has a severe impact in the efficiency of the link adaptation process and also affects the scheduling process since the upper limit of the resources allocated to each user depends on this report.

D. Hybrid HCF Scheduling Mechanism for WLAN

The classical HCF scheduling mechanism, where real time traffic is delivered through HCCA and best effort users mutually contend to get the channel control, presents a clear bottleneck for the downlink best effort users since the access point competes in the same manner with the rest of uplink users. In order to overcome this limitation, authors are also currently working using the SPHERE platform to develop a new hybrid scheme. The authors have proved that an adequate allocation of the downlink and uplink best effort traffic between EDCA and HCCA can increase the system's performance compared to the traditional scheme.

E. CRRM mechanisms for RAT selection

In terms of CRRM research, initial investigations from the authors have proposed a selection RAT algorithm, named UBReQoS algorithm [20], that intelligently distribute users among the RATs of a heterogeneous system according to the load of each RAT, the required QoS level, and the effect of a RAT selection on users already employing such RAT. In particular, the UBReQoS scheme considers three different utility functions. The first one aims at selecting the RAT that achieves the user required QoS level. A distinctive feature of this utility function is that users are not always assigned to the RAT achieving the highest transmission rate but to the RAT that ensures its QoS even if there are other available RATs with higher data rates; this procedure will enable a load balancing policy among RATs. The second utility function considers the RAT's load to avoid assigning RATs that are already overcharged which will increase transmission times and

therefore reduce user perceived QoS. The third utility function considers the effect of interfering users per RAT in order to favor the selection of RATs with lower interference levels.

Table VIII shows the performance improvements that can be achieved with the UBReQoS proposal in a multi-service scenario compared to a reference algorithm in which each service type is permanently allocated to the same RAT; the reference schemes assigns web and email users GPRS, H.263 video users with a bit rate of 32 kbps to EDGE, and H.263 video users with bit rates of 64 and 256 kbps to HSDPA (no multi-channel operation is considered here). The simulation results have been obtained using the SPHERE simulation platform.

TABLE VIII. USER SATISFACTION (%) FOR UBREQOS AND REFERENCE ALGORITHMS (SCENARIO I, SINGLE-CHANNEL OPERATION)

Service	Reference	UBReQoS	Improvement
Web	49.79	59.70	+ 9.9 %
Email	54.32	58.78	+ 4.5 %
32 kbps	88.23	91.41	+ 3.2 %
64 kbps	92.90	90.45	– 2.5 %
256 kbps	51.86	55.92	+ 4.1 %
Global	85.15	86.20	+ 1.1 %

V. CONCLUSIONS

This work has presented SPHERE, a radio simulation platform for heterogeneous wireless systems jointly developed by the University Miguel Hernández and the Polytechnic University of Valencia.

In this paper the main features of the implemented simulation tool have been briefly described. Starting from an overview of all the considered RATs, this paper has given a full account of the logical structure of SPHERE and all its modules, especially describing those related with the radio resource management strategies, which are the main concern of the research project under which the SPHERE platform is being built.

The platform has been validated through system level simulations and some initial investigations that are being conducted using SPHERE have been explained to illustrate the potential of the implemented software platform.

ACKNOWLEDGMENT

This work has been partially funded by CICYT (Spanish National Science and Technology Council) and the FEDER program of the European Commission under the project TEC2005-08211-C02-02.

REFERENCES

[1] EVEREST Project IST-2002-001858. Available at http://www.everest-ist.upc.es/.

[2] A. Tolli, P., Hakalin, and H. Holma, "Performance evaluation of common radio resource management (CRRM)". IEEE International Conference Communications (ICC'02),pp. 3429-3433, 28 April-2 May, 2002.

[3] F. Malavasi, M. Breveglieri, L. Vignali, P. Leaves and J. Huschke, "Traffic control algorithms for a multi access network scenario comprising GPRS and UMTS", 57th IEEE Vehicular Technology Conference, vol. 1, pp. 145-149, April 2003.

[4] R. Veronesi, "Multiuser scheduling with multi radio access selection", 2nd International Symposium on Wireless Communication Systems, pp. 455-459, September 2005.

[5] R. Ferrus et al., "EVEREST testbed: QoS management evaluation in B3G networks", IEEE International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities, March 2006.

[6] M. Takai et al., "Scalable Testbed for Next-Generation Wireless Networking Technologies", IEEE International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities, pp. 162-171, February 2005.

[7] 3GPP, Physical layer procedures (FDD), 3GPP TS25.214 V7.1.0.

[8] ETSI, "Universal Mobile Telecommunications System (UMTS); Selection procedures for the choice of radio transmission technologies of the UMTS", TR 101.112, Apr. 1998.

[9] Y. Wang and L. Cuthbert; "Agent-based load balancing of WLAN in indoor usage", IASTED International Conference on Wireless and Optical Communications , June 2003.

[10] J.F. Monserrat, J. Gozalvez, R. Fraile and N. Cardona, "Effect of Shadowing Correlation Modeling on the System Level Performance of Adaptive Radio Resource Management Techniques", IEEE International Symposium on Wireless Communications Systems, September 2005.

[11] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems", IEE Electronics Letters, pp. 2145–2146, November 1991.

[12] J. Gozalvez y J. Dunlop, "Link Level Modelling Techniques for Analysing the Configuration of Link Adaptation Algorithms in Mobile Radio Networks", Proceedings of European Wireless, pp 325-330, February 2004.

[13] J. Gozálvez and J.J González-Delicado, "Channel allocation mechanisms for improving QoS in packet mobile radio networks", IEE Electronics Letters, pp. 21-22, January 2005).

[14] J. Gozálvez, M. López-Benítez and O. Lázaro, "Link adaptation algorithm for improved wireless transmission of delay-sensitive packet data services", IEE Electronic Letters, pp. 813-815, July 2005.

[15] P. Barford and M. Crovella, "Generating representative web workloads for network and server performance evaluation", Joint International Conference on Measurement and Modeling of Computer Systems, June 1998.

[16] J. Ho, Y. Zhu and S. Madhavapeddy, "Throughput and buffer analysis for GSM general packet radio service (GPRS)", IEEE Wireless Communications and Networking Conference, pp. 1427-1431, September 1999.

[17] O. Lázaro, D. Girma and J. Dunlop, "H.263 video traffic modelling for low bit rate wireless communications", IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications, pp. 2124-2128, September 2004.

[18] J. Laiho, A. Wacker and T. Novosad, "Radio network planning and optimisation for UMTS", Wiley, 2002.

[19] B. Wang, K.I. Pedersen and P.E. Mogensen, "Performance of VoIP on HSDPA", IEEE Vehicular Technology Conference, pp. 2335-2339, May 2005.

[20] M. López-Benítez y J. Gozalvez, "QoS provisioning in beyond 3G heterogeneous wireless systems through common radio resource management algorithms" Proceedings of the Second ACM Workshop on QoS and Security for Wireless and Mobile Networks (Q2SWinet 2006), pp 59-66, October 2006.

EUROPEAN COOPERATION
IN THE FIELD OF SCIENTIFIC
AND TECHNICAL RESEARCH

EURO-COST

COST 2100 TD(07) 044
Lisbon, Portugal
2007/Febr/26-28

SOURCE: iTEAM, Instituto de Telecomunicaciones y Aplicaciones Multimedia,
Polytechnic University of Valencia,
Spain

SPHERE – A Simulation Tool for CRRM Investigations

José F. Monserrat, David Martín-Sacristán, David Gozálvez, Narcís Cardona
iTEAM/UPV
Edificio 8G
Camino de Vera, s/n
46022 Valencia
SPAIN
Phone: + 34-963877007 88274
Fax: + 34-963879583
Email: {jomondel,damargan,dagoser,ncardona}@iteam.upv.es

Javier Gozálvez
Signal Theory and Communications Division, UMH
Av. de la Universidad, s/n
Elche
SPAIN
Phone: + 34-966658955
Email: j.gozalvez@umh.es

SPHERE – A Simulation Tool for CRRM Investigations

***Abstract*— This paper presents SPHERE, a Simulation Platform for HEterogeneous wiREless systems, and describes its motivation, methodology and implementation approach. This advanced system level simulation platform emulates simultaneously the transmission of GPRS, EDGE Multi-slot, HSDPA and WLAN at the packet level, which allows conducting novel investigations on common radio resource management for beyond 3G systems or on the optimization of radio resource management techniques. This paper presents the simulation platform, validates it and introduces its research potential.**

I. INTRODUCTION

Strong efforts are being undertaken in the research community to define and optimize a Common Radio Resource Management (CRRM) framework [1]. The increasing complexity of current and future mobile wireless technologies requires the implementation of adequate platforms to evaluate and optimize their performance. Before considering a prototype or full-scale deployment, the use of simulation platforms is becoming increasingly common within the research community due to its cost/benefit ratio. However, it is important that to conduct meaningful and appropriate studies, such simulation platforms implement accurately the entities and process under evaluation. The implementation of such advanced simulation tools has become a very challenging task when investigating CRRM techniques since different RATs need to be simultaneously emulated in a single platform.

In this context, the University Miguel Hernández and the Polytechnic University of Valencia have developed a novel, ambitious and scalable radio simulation platform for heterogeneous wireless systems, named SPHERE (Simulation Platform for HEterogeneous wiREless systems), under a common national research project. The platform currently integrates four advanced system level simulators, emulating the GPRS (General Packet Radio Service), EDGE (Enhanced Data-rates for GSM/Global Evolution), HSDPA (High Speed Downlink Packet Access) and WLAN RATs. The unique simulation platform emulates all four RATs in parallel and at the packet level, which enables an accurate evaluation of the final user perceived QoS through the implementation of novel CRRM and RRM mechanisms. The radio interface specifications of these four technologies have been faithfully implemented in the SPHERE simulation platform, which works with a high time resolution (in the order of some milliseconds). This modeling approach validates the capability of the SPHERE simulation platform to dynamically and precisely evaluate the performance of RRM/CRRM techniques. The platform has been developed following a modular and scalable design, which guarantees an easy adaptation of the platform configuration to specific requirements, and allows the rapid integration of new RATs. In particular, the research team is considering the future expansion of the platform to also emulate the UMTS and Mobile WiMAX radio interfaces.

The interest of the SPHERE platform and the research being conducted is highlighted by the support of important companies in the mobile and wireless industries, such as Motorola, Swisscom Innovations and Telefonica I+D, to the research project developing SPHERE.

II. THE SPHERE PLATFORM

Fig. 1 shows the scenario modeled by the SPHERE platform which includes the GPRS, EDGE, HSDPA and WLAN radio interfaces. As shown in Fig. 1, the SPHERE platform does not only model the radio interface of the four technologies but also implement various RAT specific RRM features and a centralized CRRM entity. This entity directly collects specific RAT information (e.g. load, channel quality conditions, etc) and interacts with the RRM entities implemented at each RAT.

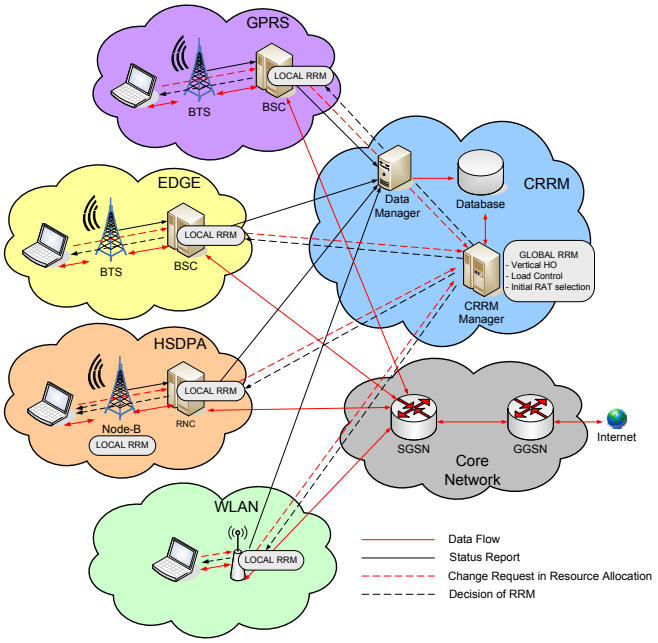

Figure 1. SPHERE heterogeneous scenario

A logical structure of the SPHERE simulation platform, which is a discrete-event system level simulator concentrated on the downlink performance, is shown in Fig. 2. The components shown in this figure, their features, interactions and data flow will be described in the following sections. Finally, the potential of the platform to conduct advanced research on the

design, evaluation and optimization of CRRM and RRM techniques will be demonstrated.

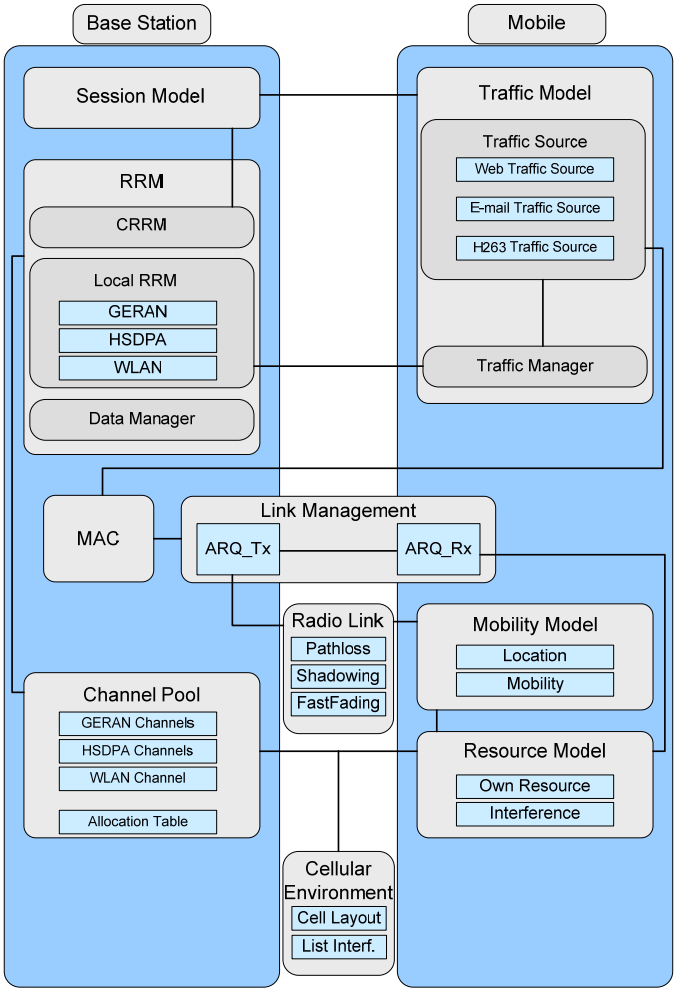

Figure 2. SPHERE logical structure

A. *Cellular Environment*

The *Cellular Environment* entity is a system module storing the location of each base station and the interfering relations among them; this information is needed to estimate the experienced interference levels. The cellular layout can be modified offline at any time to change the system configuration under study. Currently, the SPHERE platform considers a cell layout of 27 omni-directional cells. In order to avoid border effects, a wrap-around technique has been applied. Three-sectorized cell sites have also been modeled. A basic scenario has been defined considering two concentric coverage cells with increasing cell radii. The smaller one, of 50m, represents the coverage area of the WLAN technology. The second coverage radius of 500m includes HSDPA, EDGE and GPRS.

B. *Radio Link*

This module models the radio propagation conditions between transmitter and receiver and is a generic entity employed by any wireless link established. It characterizes the three radio propagation effects, namely path loss, shadowing and fast fading.

The path loss model provides an average measure of the signal attenuation over a given distance. For cellular systems, the path loss (in dB) reported in [2] has been considered:

$$L_p = \left(40 - 16 \cdot 10^{-2} \Delta h_b\right) \log_{10} d - 18 \log_{10} \Delta h_b + 21 \log_{10} f + 80 \quad (1)$$

Where d is the distance in km between the base station transceiver and the mobile terminal, f is the carrier frequency in MHz and Δh_b is the base station antenna height, measured in meters from the average roof top level. In HSDPA typical values for these parameters are Δh_b=15m and f=2000 MHz, leading to

$$L_p\left(dB\right) = 128.1 + 37.6 \log_{10} d \quad (2)$$

This model is applicable to frequencies ranging from 1.5 GHz up to 2 GHz. As a result, a carrier frequency of 1.8 GHz has been assumed for GPRS and EDGE. This change in the frequency parameter results in the following path loss equation

$$L_p\left(dB\right) = 127.2 + 37.6 \log_{10} d \quad (3)$$

For WLAN, the following path loss model has been implemented [3]:

$$L_p(dB) = 145 + 35 \log_{10} d \quad (4)$$

Various works have demonstrated the importance of an adequate shadow modeling to conduct appropriate RRM investigations [4]. In fact, the shadowing effect results in additional signal attenuation due to obstacles in the path between transmitter and receiver. Measurements have shown that the shadowing loss can be modeled as a random process with a normal distribution of mean 0 dB and standard deviation between 4 and 12 dB depending on the propagation environment. The SPHERE platform considers an urban or suburban environment and, as a result, a shadowing standard deviation equal to 6 has been set. The shadowing is a spatially correlated process so that the shadowing loss experienced by a mobile at a given position is correlated to that experienced at a nearby position. Although the authors are actually working on migrating the shadowing models to that used in [4], the current version of the SPHERE platform models this spatial correlation as detailed in [5], with a de-correlation distance of 20 m.

Fast fading modeling is also important when considering technologies such as EDGE and HSDPA, which base their radio operation on link adaptation techniques [6]. In these RATs, transmission conditions are modified depending on the current channel characteristics. For the sake of simplicity, in SPHERE only a simple block fading model is considered, i.e. the fast fading stays constant over a coherence time interval and each sample is statistically independent. Hence, in addition to the path loss and shadowing value of each radio block, in SPHERE a third multiplicative factor is considered when determining the received carrier: the fast fading coefficient. This factor is of unit mean and follows the probability density function:

$$f_\psi(\psi) = \frac{M}{(M-1)!} (M\psi)^{M-1} e^{-M\psi} \quad (5)$$

Here, *M* denotes the number of resolvable independent multi-paths at the receiver. In GPRS and EDGE *M* is set to 1 and in HSDPA *M*=3.

C. Base Station

As shown in Fig. 2, the Base Station entity is responsible for the Medium Access Control (MAC) and RRM functions. It also controls the channel pool where the status of all channels per RAT is maintained. In the Base Station the session generation process is also located. Once a new mobile is active in the system, the CRRM entity chooses its initial RAT depending on a specific policy. When a mobile station requests a channel from a given RAT, the channel pool of the serving base station is examined to search for an available channel. If a free channel is available on the requested RAT, the mobile station is assigned a randomly chosen channel or based on some quality metrics [7]. If a free channel is not available on the requested RAT, the mobile station is assigned a channel from a different RAT, depending on the CRRM scheme under consideration, or it is placed in a queue until a transmitting mobile ends its transmission and releases its channel. For users in GPRS and EDGE queues, a First-Come First-Served (FCFS) scheduling policy is applied so that channel requests are satisfied in the same order as they appear. Users in the HSDPA queue can be served either in a round robin fashion, according to the Max C/I criterion, which selects at any moment the user with better transmission quality, or following the proportional fair algorithm. In WLAN, real-time traffic is delivered through HCCA with a FCFS policy, whereas best effort users mutually contend to get the channel control being served using the EDCA protocol.

Apart from the scheduling, other implemented RRM functionalities include Link Adaptation for GPRS, EDGE, HSPDA and WLAN, multi-channel operation for GPRS and EDGE, multi-code allocation in HSDPA and call admission control in all technologies.

Link adaptation (LA), also referred to as Adaptive Modulation and Coding (AMC) in HSDPA, is an adaptive RRM technique that periodically estimates the channel quality conditions and selects the optimum transmission mode based on a predefined selection criterion. For web browsing and email services, the transmission mode that maximizes the throughput is selected. In the case of multi-channel transmissions, the channel quality conditions are estimated over all channels simultaneously assigned to a single user and their average value is used to estimate the optimum transmission mode according to the established selection criterion. In HSDPA, the mobile directly reports its channel conditions to the base station by means of the CQI. With this information, the base station knows the maximum allocable number of codes as well as the modulation and coding scheme. The final allocation shall always be as higher as possible but not exceeding the transmission mode reported by the user. For WLAN, Automatic Rate Fallback (ARF) is the implemented LA algorithm. ARF has been widely implemented in many WLAN products although they are not included in the IEEE standards. In ARF, the sender deduces the channel conditions by measuring the numbers of consecutively successful and failed transmissions. The sender adjusts its modulation mode and data rate in accordance with these measurements.

Multi-channel or multi-code mechanisms based on some innovative schemes currently under investigation have been incorporated into the SPHERE simulation platform for the GPRS-EDGE and HSDPA radio interfaces. In these schemes, the number of channels or codes that a base station can simultaneously allocate to a single user depends on factors such as the capability of the terminal, the system load, the availability of radio resource, the requested service type and the considered multi-channel allocation policy.

In the current implementation, a RAT selection can be performed for each new session, periodically, or every time a new packet is generated. It is important to highlight that RAT changes are done dynamically in the SPHERE platform and that the radio transmission can be immediately resumed with the newly selected RAT at the stage where the radio transmission ended using the previous RAT. The platform has also been prepared to consider the case in which a user handles different application sessions through various RATs.

D. User traffic behavior

User traffic demands are usually described at two levels: session-arrival process and traffic models. Session-arrival processes, also referred as traffic generation, are usually modeled as a birth-death process, which can be characterized by the following parameters: busy hour call attempts (BHCA), arrival distribution, mean session duration, duration distribution, etc. On the other hand traffic models describe the source behavior within a session. They vary depending on the type of service and they can be described by parameters such as the average active/inactive times, time distributions, data rate, packet length distribution, etc. In SPHERE, the session-arrival has been implemented at the base station, while the traffic models are controlled by the mobile station to optimize the code.

1) Session Model

Three different services have been implemented in SPHERE, namely web browsing, real-time H.263 video transmissions and email. Cellular subscribers are usually considered to have independent behavior one from each other, which results in exponentially distributed inter-session arrival times. For each one of the implemented services a specific inter-session time is defined, which allows controlling the traffic load.

2) Traffic Model

Despite considering downlink transmissions, the current version of the SPHERE platform implements the traffic models at the mobile station. This has been done to optimize the simulation code.

The web browsing service follows the model described in [8]. It follows an ON/OFF pattern where a web browsing session starts with the submission of a web page request by the user.

The email traffic model also follows an ON/OFF pattern [9]. The model assumes that incoming messages of a user are stored at a dedicated email server. This server keeps the emails in a mailbox until the user logs onto the network, following the session model previously described, and downloads the emails. When the user opens the mailbox, the headers of the available messages are downloaded. The user scans then through these headers and downloads the emails she/he is interested in. When the user finishes downloading a message

(active period), she/he will read it (inactive period) before downloading the next message, and so on.

Real-time services have also been included in SPHERE through the emulation of real-time H.263 video transmissions following the model presented in [10]. This model takes into account the three different frame types considered in the H.263 standard, namely I, P and PB. The model characterizes the size and duration of the video frames, the correlation between both parameters for each video frame, and the transition probability between different video frame types. The modeling is performed at two levels. The first one establishes the frame type to generate. I-frames are periodically created, while a Markov chain drives the transition generation between P- and PB-frames. Once the frame type is selected, the model determines the size and the duration of the video frame to be transmitted. The reader is referred to [10] for a more detailed description of the real-time H.263 video traffic model.

E. Mobile station

1) Mobility Model

The implemented mobility model considers a suburban scenario where users move at constant speed. The initial position of a mobile station within a cell is randomly selected according to a random uniform distribution. The discrete nature of event-driven simulations has been reflected in the mobility model through the definition of the time at which a mobile's movement is updated. The length of each step is constant and equal to the decorrelation distance used for the shadowing model; the position of a mobile at a particular time between two random positions is extracted by interpolation. The direction of each step is randomly established by adding a random angle to the previous direction. The random angle is obtained from a normal distribution with zero mean and a variance dependent on the mobile speed. The mobility model, leading to a long-term uniform user's density within a cell, was shown to be consistent with an analysis performed on real data provided by a mobile operator.

2) Resource Model

The resource model entity is basically responsible for controlling the radio transmission parameters of a channel currently assigned to a user and for estimating the experienced channel quality conditions, in this case carrier to interference ratio. For WLAN systems only the received power is calculated since it is the only value needed to obtain the maximum allocable bit rate.

The GPRS and EDGE CIR level is estimated as follows:

$$CIR_{GPRS/EDGE} = \frac{\frac{P_i}{L_P^{ii} \cdot L_S^{ii}} \psi_i}{\sum_{j \in \Omega} \frac{P_j}{L_P^{ij} \cdot L_S^{ij}} \psi_j + N_0 \cdot W} \quad (6)$$

where P_i is the power transmitted by the reference cell (cell i) to the user of concern, L_P^{ii} and L_S^{ii} are the path loss and shadowing loss over the link between transmitter and receiver at the reference cell, Ω is the set of active co-channel interferers, P_j is the transmission power of each one of the interfering channels, L_P^{ij} and L_S^{ij} are the path loss and shadowing loss over the link between the active transmitting interferers in cells j and the interfered receiver at the reference cell i, and $N_0 \bullet W$ represents the thermal noise at the receiver in the reference cell, with N_0 being the noise spectral density and W the bandwidth of the transmission channel. Finally ψ models the fast fading effect.

In CDMA-based systems, such as HSDPA, channelization codes for the users of the same cell are perfectly orthogonal. However, due to multi-path fading, this orthogonality decreases and some intra-cell interference component is observed. Intra-cell interference on a CDMA system is modeled by an orthogonality factor [11], which is usually denoted as α. In absence of multi-path fading, the codes are perfectly orthogonal, so $\alpha = 1$. In the worst case $\alpha = 0$, meaning that orthogonality is entirely destroyed. Typical values of α are between 0.4 and 0.9. Thus, the HSDPA CIR level can be expressed as follows:

$$CIR_{HSDPA} = \frac{\frac{P_i}{L_P^{ii} \cdot L_S^{ii}} \psi_i}{\frac{(P_{Ti} - P_i/C_i) \cdot (1-\alpha)}{L_P^{ii} \cdot L_S^{ii}} \psi_i + \sum_{j \in \Omega} \frac{P_{Tj} \psi_j}{L_P^{ij} \cdot L_S^{ij}} + N_0 \cdot W} \quad (7)$$

where P_i is the addition of the power transmitted in all the C_i channels allocated by the reference cell to the user of interest, P_{Ti} is the total power transmitted by the reference cell, Ω is the set of cells interfering the user and P_{Tj} is the total power transmitted by these interferers. In this expression, the parameters P_{Ti} and P_{Tj} also include the base station power reserved for other channels different from the HSDPA High Speed Downlink Shared Channel (HS-DSCH). The power allocated to each control channel has been extracted from [12]. ψ also models the fast fading effect.

In WLAN the received power is calculated as:

$$P_{Rx} = \frac{P_i}{L_P^{ii} \cdot L_S^{ii}} \quad (8)$$

It is worth noting that the WLAN interference from the rest of Access Points has not been considered in the SPHERE platform since the distance between WLAN transmitters is high enough to neglect it.

F. Link Management

The Link Management module handles the radio transmission and emulates channel errors.

1) Transmission Process

GPRS controls the radio transmission of RLC blocks through an Automatic Repeat reQuest (ARQ) protocol, described in specification 3GPP TS 04.60, that is implemented in SPHERE. This ARQ protocol is based on the numbering of the blocks and a selective repeat principle with sliding transmitting and receiving windows. The transmitter sends blocks and the receiver sends acknowledgment messages acknowledging correctly received blocks and requesting the retransmission of erroneously received blocks. The transmitting and receiving ARQ windows have a size of 64 RLC blocks. The reporting period, which defines how regularly the receiver sends acknowledgment messages, has been set to 16 blocks. No block losses and errors on the transmission of the acknowledgement messages have yet been considered in SPHERE. A similar ARQ protocol has been implemented for EDGE, with varying window sizes according to the number of channels simultaneously assigned to a single

user; the window sizes range vary from 64 to 1024 radio blocks. For EDGE transmissions, a 32 radio blocks has been selected.

In HSDPA, retransmission of erroneous transport blocks is performed by an N-channel stop-and-wait (SAW) ARQ protocol. In stop-and-wait schemes, the transmitter handles the transmission of a single block until it has been successfully received. In SPHERE, a maximum of 8 channels can be set up simultaneously since this is the value suggested in the standard. Block size is determined by the reported CQI. As for GPRS and EDGE, no transport block losses or errors on the acknowledgement messages have been emulated.

For WLAN, the SPHERE platform also implements an ARQ protocol. In this case only one channel SAW is employed. The transport block is a fixed-length IP packet of 1500 bytes although it is possible to perform fragmentation to improve channel utilization.

2) Channel Errors Emulation

For GPRS and EDGE, in order to decide whether a radio block is received in error, the experienced CIR is computed in the four TDMA frames used to transmit such block. After completing the transmission of a whole radio block, the four associated CIR values are averaged and a single CIR_{avg} value is obtained. This CIR_{avg} value is then mapped to a BLock Error Rate (BLER) value ($BLER_0$) by means of a Look-Up Table (LUT) such as the ones proposed in [6]. LUTs are used as a means of interfacing link and system level simulations using the link level analysis as a source of information for the system level. The link level performance is then represented by a simplified model consisting of a set of LUTs mapping the CIR to a given link quality parameter such as the BLER. Different LUTs need then to be produced for different operating conditions, e.g., transmission mode, mobile speed and propagation environments (typical urban or rural area). Once a $BLER_0$ value is obtained from the adequate LUT, a random process is used to decide whether the radio block is correctly received.

For HSDPA, the experienced CIR is computed in each slot of a TTI. After completing the transmission of a whole transport block, the three associated CIR values are averaged and a single CIR_{avg} value is obtained, which represents the quality experienced by the transport block. In the same way as in the GPRS-EDGE case, a LUT is employed to map the CIR value in a BLER value and to decide whether a block is correctly received. When a transport block is received in error, it is not discarded but stored in the receiver buffer and combined with retransmissions according to a specific method. One possibility is to use the Chase Combining (CC) scheme in which retransmitted blocks are identical to that of the first transmission. After several transmissions, the resulting effective CIR value (CIR_{eff}) is representative of the global quality experienced by the data stored in the receiver buffer after the combination process. The value of CIR_{eff} is used to decide whether the information is received in error. To this end, a random number X_0 between zero and one is drawn from a uniform distribution prior to the first transmission of a transport block. This random number is mapped to a CIR value by means of the corresponding LUT. The CIR value is then established as the minimum effective CIR (CIR_{min}) that must be obtained at the receiving side to consider that the information is correctly decoded. Each time a transport block is transmitted, the decision is taken by comparing CIR_{min} with the CIR_{eff} obtained after the combination of the current transmission and previous transmissions (if any) of the transport block. When $CIR_{eff} \geq CIR_{min}$, the transport block is assumed to be successfully received. However, if $CIR_{eff} < CIR_{min}$, the transport block is then assumed to be received in error.

When modeling WLAN it is unusual to make use of LUTs as in cellular systems. Rather, the concept of sensibility is employed. If a certain physical transmission mode is given and the mean received power is over its specific sensibility then the transmitted block will be properly received, otherwise the block is dropped.

III. RRM AND CRRM INVESTIGATIONS

Once the SPHERE simulation platform has been described, this section is devoted to validate the implemented platform and shows its potential for conducting novel CRRM and RRM investigations. In particular, this section briefly presents EDGE, HSDPA and WLAN RRM work conducted using the SPHERE platform and initial CRRM investigations considering a heterogeneous scenario made up of the GPRS, EDGE and HSDPA radio interfaces.

A. Platform Validation

The objective of this section is to validate the SPHERE platform by means of simulation results. For a rigorous validation, it would be desirable to compare the results obtained by SPHERE with the results from another source where a similar heterogeneous scenario was simulated. However, as no source for comparison has been found, this section compares the SPHERE results with the maximum possible performance of each RAT to prove the obtained results are within the expected range.

Table I shows the mean throughput performance per RAT and modeled traffic service. The EDGE RAT has been simulated considering different multi-slot configurations. These values have been obtained by simulating the diverse RATs of the system independently, i.e. not simultaneously, and considering a load of 15 users per cell in each simulation (3 for web-browsing, 3 for email, 3 for H.263 at 32 kbps, 3 for H.263 at 64 kbps, and 3 for H.263 at 256 kbps). In the simulation of WLAN, users are on the move inside the cell radius of 50m.

Results clearly show that the SPHERE platform offers a considerably wide range of transmission capabilities, highlighting its suitability for analyzing CRRM policies in a heterogeneous wireless framework. From such perspective, and given that EDGE under multi-slot operation increases the transmission bit rate ‘variety’, each EDGE multi-slot configuration could be considered as a different RAT for the CRRM policies implemented in heterogeneous systems. This results in a heterogeneous wireless framework with a larger set of radio access alternatives that offer a wider set of transmission capabilities.

TABLE I. MEAN THROUGHPUT PERFORMANCE (KBPS)

	Web	Email	H.263 32 kbps	H.263 64 kbps	H.263 256 kbps
GPRS	12.20	12.31	14.23	13.71	13.72
EDGE (1S)	22.25	21.61	22.66	21.58	26.68
EDGE (2S)	31.40	31.17	39.70	38.67	40.04
EDGE (3S)	70.20	69.45	60.85	64.57	64.03
EDGE (5S)	82.52	78.59	105.20	105.60	124.14
HSDPA	1035.59	965.318	610.425	720.58	1017.60
WLAN	18229	15285	15432	16892	17670

Although it can be observed that the performance for each service is similar for a given RAT, it is important to remember that the configuration of LA and AMC differ for real-time video, and web and email services. Despite similar performance, user satisfaction is not similarly maintained for each RAT since each service has different QoS expectations. We have then defined user satisfaction parameters per traffic service. Web-browsing and email users are assumed to be satisfied when they download a web page or an email in less than 4 seconds, as specified in 3GPP TS 22.105. Video users are supposed to be satisfied every time a video frame is entirely received before a new one is generated, i.e. no part of the video frame is discarded. The user satisfaction is therefore defined as the percentage of times that a web page, email o video frame transmission results satisfactory for the end user. Table II shows the obtained results. In general, for a given service, the user satisfaction increases as the selected RAT offers better capabilities. However, we can observe that for some H.263 users, EDGE using 5 slots obtains better user satisfaction than HSDPA despite its lower throughput performance (Table I). This is due to two main reasons. First of all, it is important to note that the defined user satisfaction parameter for video transmissions is based on the bit rate and not on the throughput since transmission errors have not been accounted in the parameter's definition. Also given the real-time nature of H.263 video transmissions, retransmissions of erroneously received data blocks for EDGE and GPRS have not been allowed. On the other hand, up to four retransmissions have been allowed for HSDPA to take profit of CC capabilities. As a result, given that the experienced BLER is quite high for the considered operating conditions, HSDPA requests several retransmissions of a transport block and therefore less video frames are transmitted before the next one is generated than considering EDGE with 5 slots.

TABLE II. USER SATISFACTION (%)

	Web	Email	H.263 32 kbps	H.263 64 kbps	H.263 256 kbps
GPRS	0.0	44.5	0.4	0.3	0.0
EDGE (1S)	53.8	53.9	87.9	43.0	0.0
EDGE (2S)	64.1	55.7	99.9	85.9	0.2
EDGE (3S)	74.4	59.7	100.0	98.9	0.6
EDGE (5S)	84.7	71.9	100.0	100.0	15.9
HSDPA	99.8	98.8	99.9	99.7	95.9
WLAN	100.0	100.0	100.0	100.0	100.0

As QoS requirements increase, RATs with higher bit-rates are required in order to obtain an acceptable degree of satisfaction. For the most demanding service, i.e. H.263 video transmission with a mean bit-rate of 256 kbps, only HSDPA and WLAN are able to offer an acceptable satisfaction level to the users for the considered operating conditions. On the other hand, for services with low QoS requirements such as background services, RATs with limited capabilities can fulfill the user expectations.

B. Effect of Channel-Quality Indicator Delay on HSDPA

The authors have also assessed the intrinsic delay of the Channel Quality Indicator (CQI) reporting process and its effect on the HSDPA system performance.

In HSDPA the user equipment (UE) is responsible for reporting periodically to the Node-B the downlink channel quality. This channel information is numerically represented by means of the CQI, which ranges from 1 up to 30, increasing its value when channel quality augments. Each CQI can be mapped into a combination of transmission parameters specifying a concrete value of the transport block size (TBS), number of simultaneous channelization codes, modulation and code rate. These parameters were specifically chosen to configure a 1dB granularity in the carrier to interference ratio (CIR) among consecutive CQIs for a Block Error Rate (BLER) level of 10%.

As the channel state information is provided by the UE to the Node-B there is a non-negligible delay between channel estimation and the reception of this information in the Node-B. Moreover, some time pass since the Node-B receives the CQI until it uses this information in the LA and scheduling mechanisms. Finally, at least 2 slots more pass until the UE begins to receive the data scheduled in the Node-B due to the constant delay between the HS-SCCH channel which signals the start of data transmission and the HS-PDSCH which handles the data. Consequently, the whole delay is around 6 ms, that is to say, three TTIs.

All of the previous delays are inherent in the HSDPA functioning and hence are not controllable. This research deals with the only delay which can be controlled by the system, i.e. the reporting period delay.

In general, the higher the reporting delay is the higher inaccurate the channel estimation. This inaccuracy depends on how the channel and therefore the CIR level vary in the time domain.

On the other hand, fast packet scheduling performed in Node B is one of the main features of HSDPA. Its implementation is not specified and investigation in this field can provide a great differentiation among different HSDPA systems. In each TTI the scheduler makes a decision about to which UEs the Node-B will transmit in the next TTI and the characteristics of this transmission (TBS, number of channelization codes, code rate and power).

The scheduling decision is made taking into account a list of candidate UEs which have data to be delivered in the Node-B buffer. The scheduler can differentiate if these data are new or are waiting for retransmissions, giving a highest priority to retransmissions. A maximum number of code multiplexed users is fixed 'a priori' in the simulations, setting a limit to the number of multiplexed UEs in each TTI.

In this paper two kinds of resource allocating schemes are analyzed. The former does not guarantee any quality of service (QoS) since it employs scheduling algorithms that are not QoS aware. The latter is based on separating different types of services what can preserve real time services from best effort services.

Within the first group some classic scheduling algorithms such as Round Robin, Max-CIR and Proportional Fair are evaluated.

Within the second group, a scheduler which serves first the RT users and later the best effort traffic has been implemented. This scheme is able to maintain the QoS for RT users if a call admission control (CAC) mechanism prevents congestion in the cell. This simple differentiation of services provides great results when considering a mixed scenario with RT and best effort users as compared with the case in which there is not differentiation.

A mixed traffic scenario is considered with 8 H.263 at 64kbps users and 8 web users. Only pedestrian users are considered.

Tables 3 and 4 collect the user satisfaction experimented by the H263 users and the web users in the simulated scenario. From these results it can be concluded that the MaxCIR algorithm presents the best performance and RR shows the worst functioning. The PF represents an intermediate point between these two algorithms since it allocates resources in a more intelligent way as compared with RR but does not take the best advantage from good channel conditions as MaxCIR does.

Due to the service differentiation there is a transfer in the user satisfaction from the less priority service to the more priority service. By means of this differentiation it is possible to obtain better results in terms of H263 user satisfaction at the expense of a poorer web user satisfaction. The results show clearly that the service differentiation becomes a good option to maintain the H263 QoS. An improvement between 5 and 10 percentage points in the H.263 user satisfaction is observed.

TABLE III. USER SATISFACTION WITHOUT SERVICE DIFFERENTIATION

	MCIR		PF		RR	
CQI reporting period (ms)	H263	WWW	H263	WWW	H263	WWW
2	0.8247	0.9715	0.8155	0.9623	0.7847	0.9540
16	0.7965	0.9693	0.7789	0.9648	0.7521	0.9512
32	0.7885	0.9654	0.7705	0.9559	0.7367	0.9513
64	0.7762	0.9672	0.7296	0.9488	0.7325	0.9434

TABLE IV. USER SATISFACTION WITH SERVICE DIFFERENTIATION

	MCIR		PF		RR	
CQI reporting period (ms)	H263	WWW	H263	WWW	H263	WWW
2	0.8741	0.9305	0.8624	0.9040	0.8464	0.8969
16	0.8749	0.9221	0.8614	0.9070	0.8423	0.8774
32	0.8570	0.9124	0.8604	0.9035	0.8320	0.8801
64	0.8636	0.9033	0.8579	0.8932	0.8211	0.8682

C. Hybrid HCF Scheduling Mechanism for WLAN

The classical HCF scheduling mechanism, where real time traffic is delivered through HCCA and best effort users mutually contend to get the channel control, presents a clear bottleneck for the downlink best effort users since the access point competes in the same manner with the rest of uplink users. In order to overcome this limitation, authors are also currently working using the SPHERE platform to develop a new hybrid scheme, called hybrid HCCA-EDCA centralized scheme (HHE-CS). Basically this algorithm transmits all RT traffic via the HCCA access protocol jointly with the best effort traffic pending for downlink transmission. Finally uplink best effort traffic is transmitted by means of contention in the EDCA period. Results have proved that this adequate allocation can increase the system's performance compared to the traditional scheme.

The behavior of HHE-CS with best effort traffic is studied analyzing the improvement in total throughput with respect to the more traditional approach. The average cell web browsing throughput obtained using HHE-CS and a scheme where all best effort traffic is transmitted through EDCA is compared. A fixed number of 25 H263 at 64kbps bidirectional users are simulated whereas the number of web users varies from 10 up to 60. The time reserved for contention free transmissions is fixed at the 90% of the time. Also, to reveal the reduction in the downlink bottleneck occasioned by contention transmissions, the ratio between the downlink web throughput and the uplink web throughput is calculated. It is worth noting that, since the traffic generated by downlink web users is in average six times higher than uplink users, this ratio is expected to be around 6. All stations are transmitting at 6 Mbps. As it can be seen in Fig. 3 and Fig. 4 both schemes have similar throughputs until the network reaches the saturation point from which the HHE-CS not only achieves an overall better throughput but also manages to maintain a higher DL/UL ratio. This is possible due to the fact that when transmitting downlink traffic through a distributed mechanism, an important part of the time is wasted during contention whereas HCCA can seize the channel and start transmitting almost immediately. Also, since downlink transmissions through HCCA are made by the AP without the contention of the uplink stations, the DL/UL ratio can be greatly improved as it is shown in the results. It can also be noted that once the number of station surpasses the saturation point, the cell performance decreases because of the increasing number of collisions that take place in the system. Although the HHE-CS achieves an important improvement in the DL/UL ratio, this is not enough to assure an adequate balance between uplink and downlink traffic. To achieve this objective, further increments of the time reserved for contention free transmissions can be made.

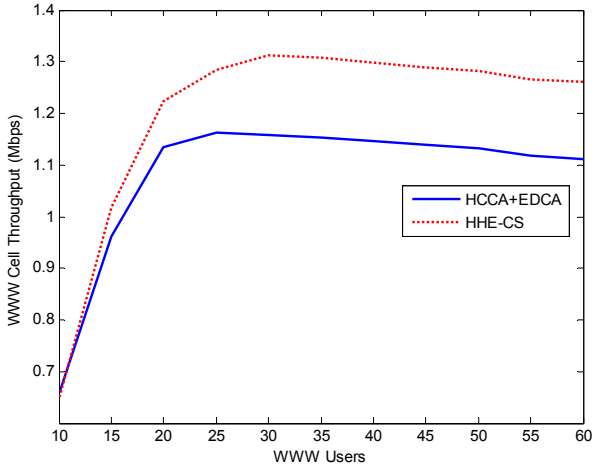

Figure 3. WWW Cell throughput for HHE-CS and HCCA+EDCA with an increasing number of WWW users

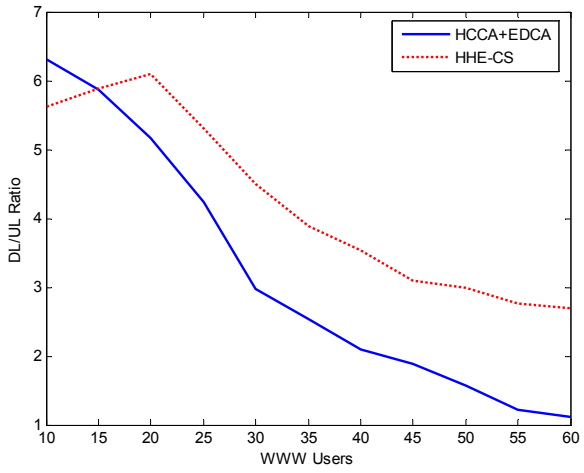

Figure 4. DL/UL ratio for HHE-CS and HCCA+EDCA with an increasing number of WWW users

D. CRRM mechanisms for RAT selection

Regarding CRRM research, initial investigations from the authors have proposed a selection RAT algorithm, named UBReQoS algorithm [13], that intelligently distribute users among the RATs of a heterogeneous system according to the load of each RAT, the required QoS level, and the effect of a RAT selection on users already employing such RAT. In particular, the UBReQoS scheme considers three different utility functions. The first one aims at selecting the RAT that achieves the user required QoS level. A distinctive feature of this utility function is that users are not always assigned to the RAT achieving the highest transmission rate but to the RAT that ensures its QoS even if there are other available RATs with higher data rates; this procedure will enable a load balancing policy among RATs. The second utility function considers the RAT's load to avoid assigning RATs that are already overcharged which will increase transmission times and therefore reduced user perceived QoS. The third utility function considers the effect of interfering users per RAT in order to favor the selection of RATs with lower interference levels.

Table V shows the performance improvements that can be achieved with the UBReQoS proposal in a multi-service scenario compared to a reference algorithm in which each service type is permanently allocated to the same RAT; the reference schemes assigns web and email users GPRS, H.263 video users with a bit rate of 32 kbps to EDGE, and H.263 video users with bit rates of 64 and 256 kbps to HSDPA (no multi-channel operation is considered here). The simulation results have been obtained using the SPHERE simulation platform.

TABLE V. USER SATISFACTION (%) FOR UBREQOS AND REFERENCE ALGORITHMS (FIRST TILITY FUNCTION, SINGLE-CHANNEL OPERATION)

Service	Reference	UBReQoS	Improvement
Web	49.79	59.70	+ 9.9 %
Email	54.32	58.78	+ 4.5 %
32 kbps	88.23	91.41	+ 3.2 %
64 kbps	92.90	90.45	– 2.5 %
256 kbps	51.86	55.92	+ 4.1 %
Global	85.15	86.20	+ 1.1 %

IV. CONCLUSIONS

This work has presented SPHERE, a radio simulation platform for heterogeneous wireless systems jointly developed by the University Miguel Hernández and the Polytechnic University of Valencia.

In this paper the main features of the implemented simulation tool have been briefly described. This paper has given a full account of the logical structure of SPHERE and all its modules, especially describing those related with the radio resource management strategies, which are the main concern of the research project under which the SPHERE platform is being built.

The platform has been validated through system level simulations and some initial investigations that are being conducted using SPHERE have been explained to illustrate the potential of the implemented software platform.

ACKNOWLEDGMENT

This work has been partially funded by CICYT (Spanish National Science and Technology Council) and the FEDER program of the European Commission under the project TEC2005-08211-C02.

REFERENCES

[1] EVEREST Project IST-2002-001858. Available at http://www.everest-ist.upc.es/.

[2] ETSI, "Universal Mobile Telecommunications System (UMTS); Selection procedures for the choice of radio transmission technologies of the UMTS", TR 101.112, Apr. 1998.

[3] Y. Wang and L. Cuthbert; "Agent-based load balancing of WLAN in indoor usage", IASTED International Conference on Wireless and Optical Communications, June 2003.

[4] J.F. Monserrat, J. Gozalvez, R. Fraile and N. Cardona, "Effect of Shadowing Correlation Modeling on the System Level Performance of Adaptive Radio Resource Management Techniques", IEEE ISWCS, September 2005.

[5] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems", IEE Electronics Letters, pp. 2145–2146, November 1991.

[6] J. Gozalvez y J. Dunlop, "Link Level Modelling Techniques for Analysing the Configuration of Link Adaptation Algorithms in Mobile Radio Networks", Proceedings of European Wireless, pp 325-330, February 2004.

[7] J. Gozálvez and J.J González-Delicado, "Channel allocation mechanisms for improving QoS in packet mobile radio networks", IEE Electronics Letters, pp. 21-22, January 2005.

[8] P. Barford and M. Crovella, "Generating representative web workloads for network and server performance evaluation", Joint International Conference on Measurement and Modeling of Computer Systems, June 1998.

[9] J. Ho, Y. Zhu and S. Madhavapeddy, "Throughput and buffer analysis for GSM general packet radio service (GPRS)", IEEE Wireless Communications and Networking Conference, pp. 1427-1431, September 1999.

[10] O. Lázaro, D. Girma and J. Dunlop, "H.263 video traffic modelling for low bit rate wireless communications", IEEE PIMRC, pp. 2124-2128, September 2004.

[11] J. Laiho, A. Wacker and T. Novosad, "Radio network planning and optimisation for UMTS", Wiley, 2002.

[12] B. Wang, K.I. Pedersen and P.E. Mogensen, "Performance of VoIP on HSDPA", IEEE Vehicular Technology Conference, pp. 2335-2339, May 2005.

[13] M. López-Benítez y J. Gozalvez, "QoS provisioning in beyond 3G heterogeneous wireless systems through common radio resource management algorithms" Proc. Q2SWinet, pp 59-66, October 2006.