UNIVERSIDAD MIGUEL HERNÁNDEZ DE ELCHE

ESCUELA POLITÉCNICA SUPERIOR DE ELCHE

INGENIERÍA DE TELECOMUNICACIÓN

"ALGORITMOS DE SELECCIÓN DE TECNOLOGÍA DE ACCESO RADIO EN REDES HETEROGÉNEAS DE COMUNICACIONES MÓVILES DE CUARTA GENERACIÓN"

PROYECTO FIN DE CARRERA

Julio - 2006

AUTOR: Miguel López Benítez
DIRECTOR: Javier Gozálvez Sempere

VISTO BUENO Y CALIFICACIÓN DEL PROYECTO

Título proyecto:

Algoritmos de selección de tecnología de acceso radio en redes heterogéneas de comunicaciones móviles de cuarta generación

Proyectante: Miguel López Benítez

Director/es: Javier Gozálvez Sempere

VºBº director/es del proyecto: Fdo.: Javier Gozálvez Sempere Fdo.:

Lugar y fecha: _____

CALIFICACIÓN NUMÉRICA ☐ MATRÍCULA DE HONOR ☐

Méritos justificativos en el caso de conceder Matrícula de Honor:

_____

_____

_____

_____

Conforme presidente:	Conforme secretario:	Conforme vocal:
Fdo.:	Fdo.:	Fdo.:

Lugar y fecha: _____

AGRADECIMIENTOS

Deseo expresar mi gratitud hacia el Dr. Javier Gozálvez Sempere, director del presente trabajo, por los numerosos y valiosos conocimientos transmitidos desde que hace ya algunos años me introdujera en el campo de la investigación en comunicaciones móviles. Agradezco así mismo sus sabios consejos y sugerencias, que sin duda alguna han contribuido a mejorar el contenido de este trabajo. Agradezco también la concesión de una beca de investigación para la realización de este trabajo al Centro de Investigación Operativa de la Universidad Miguel Hernández a través del proyecto de investigación "Diseño e implementación de técnicas de decisión multi-agente aplicadas a la gestión eficiente de los recursos radio en sistemas de comunicaciones móviles" (GV05/189) financiado por la Consellería de Empresa, Universidad y Ciencia de la Generalitat Valenciana. Por último, aunque no por ello menos importante, deseo agradecer a mi familia, en especial a mis padres, y también a Yolanda, su constante e incondicional apoyo; a ellos está dedicado este trabajo.

PUBLICACIONES

El presente trabajo ha dado lugar hasta el momento a las siguientes publicaciones:

- Miguel López-Benítez, Javier Gozálvez, "QoS Provisioning in Beyond 3G Heterogeneous Wireless Systems Through Common Radio Resource Management Algorithms", *2nd ACM Workshop on QoS and Security for Wireless and Mobile Networks*, Torremolinos, 2-6 Octubre 2006, aceptado para su publicación.
- Miguel López-Benítez, María del Carmen Lucas-Estañ, Javier Gozálvez, "A Dynamic Radio Simulation Platform for the Study of Radio Resource Management Techniques in Heterogeneous Wireless Systems", *The 9-th ACM/IEEE International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, Torremolinos, 2-6 Octubre 2006, aceptado para su publicación.
- Miguel López Benítez, Javier Gozálvez, María del Carmen Lucas Estañ, "SPHERE - Plataforma Avanzada de Simulación para Sistemas Inalámbricos Heterogéneos", *XXI Simpósium Nacional de la Unión Científica Internacional de Radio*, Oviedo, 12-15 Septiembre 2006, aceptado para su publicación.

ÍNDICE GENERAL

Lista de acrónimos **21**

1. Introducción **25**

2. Trabajo Previo Relacionado **35**

2.1. Revisión de la literatura 36

2.2. Proyectos de investigación y estandarización 49

2.3. Sumario 50

3. Algoritmos Propuestos **53**

3.1. Algoritmo UBET 54

3.2. Algoritmo UBReQoS 62

3.3. Sumario 72

4. Modelos y Herramientas de Simulación **75**

4.1. Simulación a nivel de enlace y de sistema 76

4.2. Plataforma de simulación 77

4.2.1. Entorno celular 77

4.2.2. Enlace radio 79
4.2.3. Estación base 81
4.2.4. Terminal móvil 82
4.2.4.1. Modelo de movilidad 82
4.2.4.2. Estimación de interferencia 83
4.2.4.3. Modelo de tráfico 84
Modelo de tráfico de navegación web 85
Modelo de tráfico de correo electrónico 85
Modelo de tráfico de vídeo H.263 88
4.2.4.4. Módulo CRRM 91
4.2.5. Control de enlace 92
4.3. Configuración 97
4.4. Sumario 100

5. Rendimiento de los Algoritmos Propuestos **103**
5.1. Parámetros y escenarios de evaluación 104
5.2. Rendimiento del algoritmo UBET 106
5.3. Rendimiento del algoritmo UBReQoS 115
5.4. Método para estimar la calidad radio 134
5.5. Sumario 140

6. Conclusiones **143**

A. Características de las Tecnologías Consideradas **147**
A.1. Características de GPRS 148
A.1.1. Acceso múltiple en GPRS 148
A.1.2. Modos de transmisión en GPRS 149

A.1.3. Transmisión de datos en GPRS 152
A.1.4. Protocolo de retransmisión en GPRS 152
A.2. Características de EDGE 156
A.2.1. Acceso múltiple en EDGE 156
A.2.2. Modos de transmisión en EDGE 156
A.2.3. Transmisión de datos en EDGE 159
A.2.4. Protocolo de retransmisión en EDGE 159
A.3. Características de HSDPA 160
A.3.1. Acceso múltiple en HSDPA 160
A.3.2. Modos de transmisión en HSDPA 163
A.3.3. Transmisión de datos en HSDPA 168
A.3.4. Protocolo de retransmisión en HSDPA 169

Bibliografía **190**

ÍNDICE DE FIGURAS

1.1. Líneas de evolución tecnológica de los sistemas móviles celulares seguidas por operadores de todo el mundo 26

1.2. Concepto de red heterogénea 29

1.3. Ejemplo de cobertura jerárquica en una red heterogénea 33

2.1. Ejemplo de algoritmo basado en asignaciones preestablecidas 38

3.1. Función de utilidad U_1 del algoritmo UBET para GPRS 55

3.2. Función de utilidad U_1 del algoritmo UBET para EDGE 56

3.3. Función de utilidad U_1 del algoritmo UBET para HSDPA 56

3.4. Función de utilidad U_2 del algoritmo UBET 58

3.5. Aproximación para estimar el nivel de señal/portadora (en GPRS, EDGE y HSDPA) y de interferencia intracelda (sólo en HSDPA) 60

3.6. Aproximación para estimar el nivel de interferencia intercelda 60

3.7. Función de utilidad u_1 del algoritmo UBReQoS para un servicio con $R_{nom} = 12$ kbps 64

3.8. Función de utilidad u_1 del algoritmo UBReQoS para distintos servicios con CUN = 1.0 64

3.9. CDF de los tamaños de página web 71

3.10. CDF de los tamaños de correo electrónico 71

4.1. Estructura lógica de la plataforma de simulación 78
4.2. Estructura celular de la plataforma de simulación 79
4.3. Movimiento de un terminal móvil dentro de una celda 82
4.4. Modelo de tráfico de navegación web 86
4.5. Modelo de tráfico de correo electrónico 87
4.6. Codificación de vídeo H.263 90
4.7. Modelo de Markov para la transición entre tramas H.263 90
4.8. Modo de utilización de las *Look-Up Tables* (LUTs) 94
4.9. Modelado de la calidad del canal en GPRS y EDGE 95
4.10. Modelado de la calidad del canal en HSDPA 95

5.1. Porcentaje de selección global de cada RAT para el algoritmo UBET . 107
5.2. Número medio de usuarios en cola para el algoritmo UBET 112
5.3. Satisfacción global de los usuarios para el algoritmo UBET 113
5.4. Número de cambios de RAT por usuario y por minuto para el algoritmo UBET 113
5.5. Porcentaje de selección de cada RAT para los algoritmos UBET y UBReQoS 116
5.6. CDF del tiempo de transmisión de páginas web 121
5.7. Nivel de satisfacción de cada servicio con los algoritmos UBReQoS y de asignaciones fijas 126
5.8. Porcentaje de selección de cada RAT en función del valor del parámetro CUN 129
5.9. Porcentaje de selección de cada RAT en función del valor del parámetro Y 129
5.10. CDF del error de CIR/SIR cometido por el método propuesto 136
5.11. Valor medio de la relación CIR en función de la posición del terminal móvil 138

A.1. Acceso múltiple en GPRS 149
A.2. BLER versus CIR para los modos de transmisión de GPRS 151
A.3. *Throughput* versus CIR para los modos de transmisión de GPRS 151
A.4. Transmisión de datos en GPRS 153
A.5. Retransmisión de datos en GPRS 155
A.6. Familias de modos de transmisión de EDGE 157
A.7. BLER versus CIR para los modos de transmisión de EDGE 158
A.8. *Throughput* versus CIR para los modos de transmisión de EDGE 158
A.9. Acceso múltiple en HSDPA 162
A.10. BLER versus SIR para los modos de transmisión de HSDPA 164
A.11. *Throughput* versus SIR para los modos de transmisión de HSDPA . . . 164
A.12. Diferentes tipos de esquemas HARQ 167
A.13. Retransmisión de datos en HSDPA 170

ÍNDICE DE TABLAS

1.1. Principales parámetros de diferentes tecnologías de acceso radio 28

3.1. Estadísticas de los tamaños de página web 70

3.2. Estadísticas de los tamaños de correo electrónico 70

4.1. Valor de los parámetros A y B para el modelo Okumura-Hata COST 231 80

4.2. Parámetros representativos del modelo de tráfico de navegación web . . 86

4.3. Parámetros representativos del modelo de tráfico de correo electrónico . 87

4.4. Parámetros QPSK (3 km/h) 94

4.5. Parámetros 16-QAM (3 km/h) 94

4.6. Configuración de la plataforma de simulación 98

5.1. Escenarios de evaluación 105

5.2. Satisfacción global de los usuarios con el algoritmo UBET 109

5.3. *Throughput* para cada servicio con los algoritmos UBET y UBReQoS . 118

5.4. Nivel de satisfacción de cada servicio con los algoritmos UBET y UBReQoS 120

5.5. *Throughput* de cada servicio con los algoritmos UBReQoS y de asignaciones fijas 125

5.6. Número de cambios de RAT por minuto y usuario para el algoritmo UBReQoS 125

5.7. Nivel de satisfacción de cada servicio con el algoritmo UBReQoS para varias combinaciones de los parámetros de diseño del algoritmo 131

5.8. Rendimiento del algoritmo UBReQoS con la configuración B 133

5.9. Nivel de satisfacción de cada servicio con los algoritmos UBET y UBReQoS 133

5.10. Nivel de satisfacción de cada servicio 137

A.1. Modos de transmisión de GPRS 150

A.2. Modos de transmisión de EDGE 157

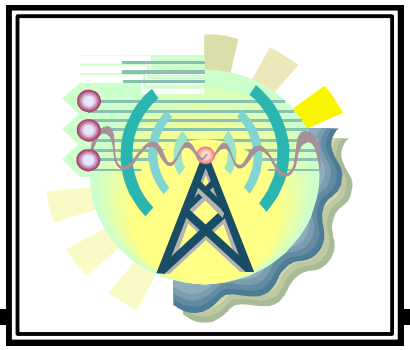
A.3. Modos de transmisión de HSDPA 163

LISTA DE ACRÓNIMOS

1G	Primera Generación
2G	Segunda Generación
2.5G	Generación 2.5
2.75G	Generación 2.75
3G	Tercera Generación
3.5G	Generación 3.5
3GPP	3rd Generation Partnership Project
4G	Cuarta Generación
AMC	Adaptive Modulation and Coding
AMPS	Advanced Mobile Phone Service
ARQ	Automatic Repeat reQuest
B3G	Beyond 3G
BEC	Backward Error Correction
BLER	BLock Error Rate
BSN	Block Sequence Number
CC	Chase Combining
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CIR	Carrier-to-Interference Ratio
CRRM	Common Radio Resource Management
CS	Coding Scheme
CSE	Circuit Switched Equivalent
CUN	Coeficiente de Utilidad Nominal
DAB	Digital Audio Broadcasting
DBS	Digital Broadcasting System
DECT	Digital Enhanced Cordless Telecommunications
DS	Direct Sequence
DSA	Dynamic Spectrum Allocation

DS-CDMA	Direct-Sequence Code Division Multiple Access
DSCH	Downlink Shared CHannel
DVB	Digital Video Broadcasting
DVB-T	Digital Video Broadcasting - Terrestrial
EDGE	Enhanced Data-rates for GSM/Global Evolution
ETSI	European Telecommunications Standardization Institute
FCFS	First-Come First-Served
FCS	Fast Cell Selection
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FH	Frequency Hopping
FIFO	First-In First-Out
GERAN	GSM/EDGE Radio Access Network
GMSK	Gaussian Minimum Shift Keying
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile communications
HARQ	Hybrid Automatic Repeat reQuest
HSCSD	High Speed Circuit Switched Data
HSDPA	High Speed Downlink Packet Access
HS-DSCH	High Speed Downlink Shared CHannel
HIPERLAN/2	HIgh PErformance Radio Local Area Network Type 2
HIPERMAN	HIgh PErformance Radio Metropolitan Area Network
HTTP	HyperText Transfer Protocol
IR	Incremental Redundancy
ISO	International Standardization Organization
IST	Information Society Technologies
ITU	International Telecommunication Union
kbps	Kilobit por segundo
LA	Link Adaptation
MAC	Medium Access Control
Mbps	Megabit por segundo
MBS	Mobile Broadband Systems
Mcps	Megachip por segundo
MCS	Modulation and Coding Scheme
MIMO	Multiple Input Multiple Output
NMT	Nordic Mobile Telephone
OFDM	Orthogonal Frequency Division Duplex
OSI	Open System Interconnection
PDC	Personal Digital Cellular
PHS	Personal Handyphone System
PSK	Phase Shift Keying
QoS	Quality of Service
RA	Radio Access

RAN	Radio Access Network
RAT	Radio Access Technology
RLC	Radio Link Control
RRM	Radio Resource Management
SAW	Stop-and-Wait
SIR	Signal-to-Interference Ratio
SNR	Signal-to-Noise Ratio
SNS	Sequence Number Space
SOHO	Small Office Home Office
TACS	Total Access Communication System
TB	Transport Block
TBS	Transport Block Set
TCP	Transmission Control Protocol
TD-CDMA	Time Division - Code Division Multiple Access
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TD-SCDMA	Time Division - Synchronous Code Division Multiple Access
TM	Transmission Mode
TTI	Transmission Time Interval
UBET	Utility Based on Expected Throughput
UBReQoS	Utility Based on Required Quality of Service
UMTS	Universal Mobile Telecommunications System
UTRAN	UMTS Terrestrial Radio Access Network
WCDMA	Wideband Code Division Multiple Access
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WMAN	Wireless Metropolitan Area Network
WPAN	Wireless Personal Area Network
WS	Window Size



CAPÍTULO 1

Introducción

Desde su lanzamiento comercial, que podría identificarse con la introducción de los sistemas analógicos de primera generación NMT (*Nordic Mobile Telephone*) y AMPS (*Advanced Mobile Phone Service*) a finales de los años 70 y comienzos de los años 80 y del sistema TACS (*Total Access Communication System*) poco después [1], la evolución de los sistemas celulares vino determinada por la adopción de la tecnología digital y el sorprendente aumento observado en el número de usuarios de la mano del sistema de segunda generación GSM (*Global System for Mobile communications*) [2]. Con la implantación de este estándar, primero europeo y luego mundial, el mercado de la telefonía móvil experimentó un crecimiento espectacular que superó todas las previsiones, convirtiéndose en una economía de escala con millones de consumidores en todo el mundo. El elevado éxito de los sistemas de comunicaciones móviles de segunda generación (2G), diseñados para proveer principalmente servicios de voz y algunos servicios complementarios de datos de baja velocidad, junto con el amplio abanico de posibilidades que ofrecía Internet (cuyo desarrollo fue también espectacular) para la prestación de nuevos servicios de banda ancha, motivaron el desarrollo hacia los sistemas de tercera generación (3G). El objetivo era lograr la integración de los sistemas móviles celulares e Internet para así poder proveer servicios de datos de alta velocidad de forma ubicua. Durante el proceso de evolución desde la 2G hacia la 3G surgieron varias tecnologías celulares [3–5]. En Europa, por ejemplo, surgieron HSCSD (*High Speed Circuit Switched Data*) [6–8], GPRS (*General Packet Radio Service*) [9–11] a veces denominado sistema 2.5G, o EDGE (*Enhanced Data-rates for GSM/Global Evolution*) [12–14] a veces denominado sistema 2.75G, hasta llegar al actual sistema europeo de tercera generación UMTS (*Universal Mobile Telecommunications System*) [15, 16].

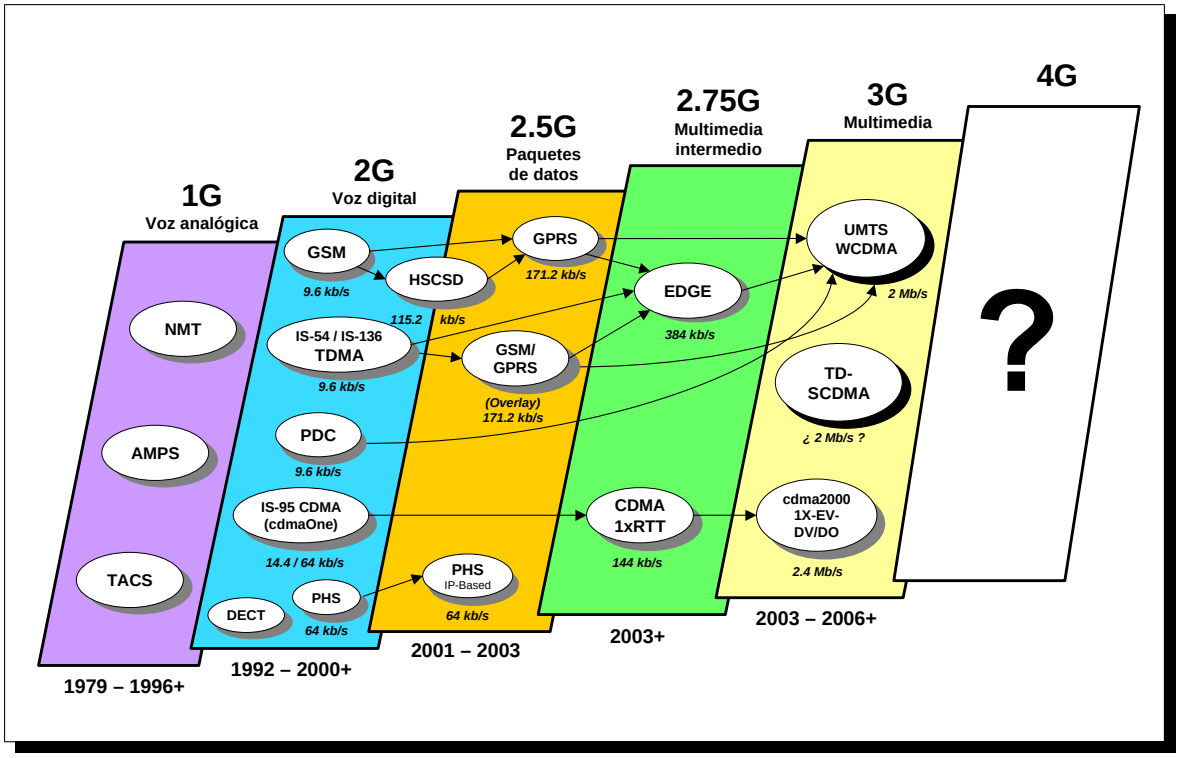


Figura 1.1: Líneas de evolución tecnológica de los sistemas móviles celulares seguidas por operadores de todo el mundo (consultar la lista de acrónimos).

La figura 1.1 muestra la evolución de los sistemas celulares en Europa, Estados Unidos y Japón (sistemas que posteriormente se han extendido al resto del mundo). Cada sistema posee sus propias características técnicas e introduce diferentes modificaciones respecto a sistemas anteriores. No obstante, existe una característica común en cada nuevo paso de la evolución, y ésta es la introducción de modificaciones en las especificaciones de la interfaz radio para conseguir tasas binarias de transmisión cada vez más elevadas. Los valores indicados en la figura 1.1 corresponden a la tasa binaria máxima teórica que puede obtenerse en la interfaz radio de cada sistema.

De forma paralela al proceso de evolución 2G/3G de los sistemas celulares, se desarrollaron numerosas tecnologías radio como, por ejemplo, las redes inalámbricas de área personal (*Wireless Personal Area Networks*, WPANs), en las que se engloba la familia de estándares IEEE 802.15 [17–19] basada en Bluetooth [20], las redes inalámbricas de área local (*Wireless Local Area Networks*, WLANs), en las que se engloba la familia de estándares IEEE 802.11 [18, 19, 21] o HIPERLAN/2 (*HIgh PErformance Radio Local Area Network Type 2*) [22–24], las redes inalámbricas de área metropolitana (*Wireless Metropolitan Area Networks*, WMANs), en las que se engloba la familia de estándares IEEE 802.16 [19, 25] o HIPERMAN (*HIgh PErformance Radio Metropolitan Area Network*) [26] impulsados por el forum WiMAX (*Worldwide Interoperability for Microwave Access*) [27], o los sistemas digitales de difusión (*Digital Broadcasting Systems*, DBSs) tales como DVB (*Digital Video Broadcasting*) [28, 29] o DAB (*Digital Audio Broadcasting*) [30] entre otras muchas tecnologías de acceso radio [31–33].

Los sistemas anteriores fueron concebidos y diseñados de forma independiente, basados en tecnologías de transmisión diferentes y orientados hacia diversos tipos de usuarios y servicios. Por ello, cada sistema posee diferentes características de velocidad de transmisión, cobertura, capacidad, coste, calidad de servicio o movilidad de los usuarios, tal y como se muestra de forma ilustrativa en la tabla 1.1. En general, cada sistema posee sus propias ventajas e inconvenientes y ninguno de ellos puede, de forma individual, dar soporte a todos los posibles tipos de servicio y requisitos de usuario existentes actualmente en el mercado de las comunicaciones inalámbricas.

Mientras la industria de las comunicaciones móviles actualmente dirige sus esfuerzos hacia el desarrollo y explotación de los recientes sistemas 3G dentro del amplio abanico de tecnologías radio existente, investigadores de todo el mundo han comenzado a plantearse en los últimos años cómo serán los sistemas posteriores a la 3G (*Beyond 3G*, B3G) o sistemas de cuarta generación (4G), con el fin de identificar las cuestiones de mayor relevancia y establecer las líneas de investigación hacia dichos sistemas [34–40].

Inicialmente, algunos investigadores concibieron los futuros sistemas de comunicaciones móviles siguiendo el mismo paradigma aplicado en el diseño de los sistemas de generaciones anteriores, es decir, buscando nuevas bandas de frecuencia para un nuevo estándar mundial, el cual habría de estar caracterizado por una nueva tecnología radio que permitiera obtener mayores velocidades de transmisión que la generación anterior [41]. Este enfoque, que podría denominarse *visión lineal*, concebía los sistemas 4G como una extensión de los sistemas anteriores y seguía poniendo el énfasis en la obtención de mayores velocidades de transmisión en la interfaz radio. Este fue, por ejemplo, el caso del concepto MAGIC (*Mobile multimedia; Anytime, anywhere, anyone; Global mobility support; Integrated wireless solution; Customized personal service*) introducido por el operador japonés NTT DoCoMo [42].

No obstante, si bien es cierto que se prevé la necesidad de mayores tasas binarias de transmisión en la interfaz radio (del orden de 100 Mbps en entornos de alta movilidad y hasta 1 Gbps en entornos de movilidad más reducida [36]), ésta no parece ser la única meta propuesta para los futuros sistemas de comunicaciones móviles. La Comisión Europea, por ejemplo, ha señalado [43] que los sistemas 4G habrán de ser capaces de proveer servicios a través de una amplia multitud de sistemas inalámbricos, incluyendo sistemas tanto públicos como privados y entornos tantos interiores como exteriores. Además, varios autores argumentan [43] que el propio usuario será el foco de atención de los futuros sistemas de comunicaciones móviles, hecho que forzará una inevitable integración de los actuales sistemas radio, los cuales habrán de ser capaces de funcionar de forma cooperante y totalmente transparente para el usuario.

Actualmente la comunidad investigadora parece coincidir ampliamente en señalar que la futura generación de comunicaciones móviles no consistirá en una única nueva tecnología radio de acceso celular con prestaciones mejoradas, sino en la integración y gestión conjunta de los actuales sistemas celulares 2G/3G, no sólo entre sí, sino también con otras tecnologías radio complementarias ya existentes y ampliamente implantadas como, por ejemplo, WLAN [44–56], sistemas de difusión DVB [57, 58] y DAB [59], o incluso sistemas de comunicaciones vía satélite [31–33] así como otras posibles futuras

Sistema	Velocidad	Tecnología	Cobertura	Movilidad	Frecuencias	Aplicación
GSM incluyendo HSCSD, GPRS y EDGE	9.6 - 384 kb/s	TDMA/FDMA, FDD	35 km para GSM, menor para datos	Alta	900, 1800, 1900 MHz	Entornos públicos y privados
UMTS	< 2 Mb/s	Familia IMT-2000, WCDMA (FDD) + TD-CDMA (TDD)	30 m - 20 km	Alta	2 GHz	Entornos públicos y privados
DECT	< 2 Mb/s	TDMA, TDD	< 50 m	Baja	1880 - 1900 MHz	Entornos de oficina y residenciales
Bluetooth	< 721 kb/s	DS (*Direct Sequence*) o FH (*Frequency Hopping*)	0.1 - 10 m	Muy baja	Banda ISM 2.4 GHz	Sustitución de cables de conexión y SOHO
HIPERLAN/2	< 54 Mb/s	OFDM, TDD	50 - 300 m	Baja	5 GHz	Entornos corporativos y *hot spots* públicos
IEEE 802.11a	< 54 Mb/s	OFDM, TDD	50 - 300 m	Baja	5 GHz	Entornos corporativos y *hot spots* públicos
DAB	1.5 Mb/s	OFDM	< 100 km	Alta	176 - 230 MHz 1452 - 1467.5 MHz	Difusión de audio
DVB-T	5 - 31 Mb/s por canal de 8 MHz	OFDM	< 100 km	Media - alta	Bandas TV debajo de 860 MHz	Difusión de vídeo

Tabla 1.1: Principales parámetros de diferentes tecnologías de acceso radio.

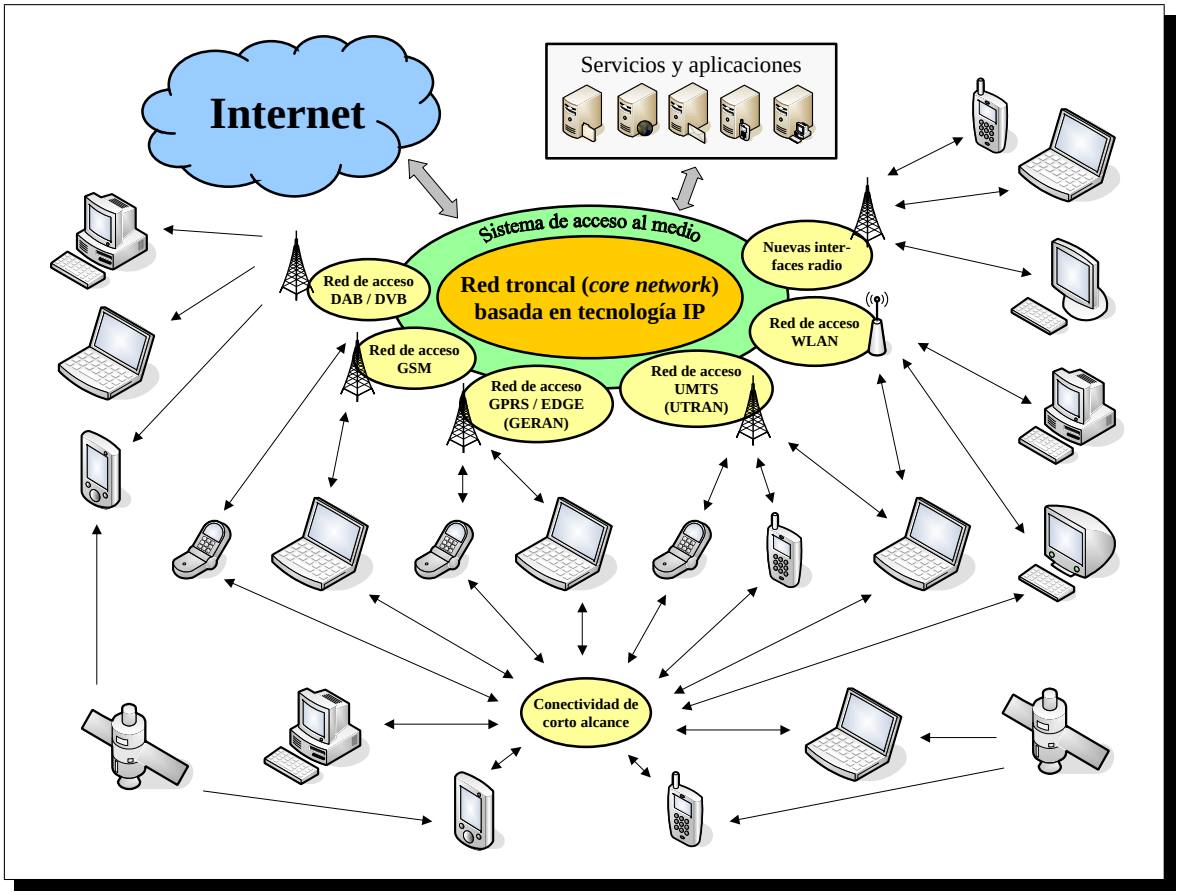


Figura 1.2: Concepto de red heterogénea [5].

tecnologías, donde una plataforma común basada en tecnología IP será el factor integrador [60–63] (ver figura 1.2). Esta *visión concurrente*, que enfatiza la heterogeneidad y la integración de redes y servicios, se basa en la idea de que las diversas tecnologías radio existentes pueden ser componentes complementarios y cooperantes de una estructura radio heterogénea, que permita el acceso a una red IP y a través de la cual los operadores puedan ofrecer acceso inalámbrico a avanzados servicios multimedia de banda ancha de forma eficiente, tanto en términos de coste como de calidad de servicio.

Esta integración no debe entenderse como un proceso de unificación y reemplazo progresivo de los sistemas de comunicaciones radio existentes ni como un obstáculo para la aparición de otros nuevos. De hecho, con un alto grado de probabilidad, la 4G no va a frenar en modo alguno la evolución de los estándares existentes ni tampoco va a dificultar la aparición de nuevos sistemas que cubran nuevos nichos de comunicación no cubiertos por los sistemas actuales. Así pues, en el contexto actual de evolución tecnológica y de mercado en comunicaciones móviles, se prevé que las redes de la siguiente generación estarán abocadas a gestionar un amplio y variado conjunto de servicios multimedia [64], cada uno de ellos con diversos requisitos de calidad de servicio, bajo la presencia de diversas tecnologías radio y terminales de usuario multi-estándar [65] capaces de reconfigurarse [66–68] para adaptarse a las características particulares de diferentes tecnologías radio, dando lugar a lo que se denominan *redes heterogéneas*.

El interés de estas infraestructuras radio heterogéneas es múltiple, pues aportará beneficios tanto a los usuarios como a los proveedores de servicios y los proveedores de red [69, 70]. Desde el punto de vista del proveedor de red, cuando una determinada red está saturada puede resultar preferible dirigir nuevos usuarios hacia otra red radio alternativa (que puede pertenecer al mismo proveedor o a un proveedor diferente asociado) en lugar de rechazar dichos usuarios o tener que aceptarlos dentro de la red saturada empeorando el nivel global de calidad de servicio del sistema. En general, el concepto de red heterogénea permitirá a los proveedores de red realizar un mejor dimensionamiento de la red y afrontar de forma más efectiva situaciones problemáticas tales como los aumentos de tráfico en ciertas horas del día (variación temporal de la demanda) o en los denominados *hot spots* o zonas de elevada densidad de tráfico (variación espacial de la demanda), dando como resultado una mejor explotación de la infraestructura de red. Desde el punto de vista de los usuarios y de los proveedores de servicios, la tecnología radio empleada para proveer el servicio resulta irrelevante mientras los requisitos de coste y calidad de servicio establecidos se vean satisfechos. Además, los intereses tanto de los usuarios como de los proveedores de servicios pueden verse beneficiados por la aparición de nuevos y sofisticados servicios. Supongamos, a modo de ejemplo, un usuario que desea conocer qué películas se ofrecen en los cines más cercanos a su posición actual. El terminal de usuario podría hacer uso de la señal enviada por los satélites de posicionamiento global GPS (*Global Positioning System*) para conocer la posición actual. Una vez determinada la posición, enviaría un pequeño mensaje de solicitud para consultar los cines próximos a esa posición que, en principio, podría enviarse a través de una red con velocidades de transmisión relativamente modestas como GPRS o EDGE. A través de una red con mayores velocidades de transmisión como, por ejemplo, UMTS, podría recibir una lista de los cines más cercanos con las películas ofrecidas en cartelera acompañada de los *trailers* de cada una de ellas. El usuario podría volver a enviar una solicitud de reserva o compra de entradas para la película a través de GPRS o EDGE o, si el aforo del cine está completo, podría adquirir la película para poder verla en su propio terminal móvil a través de una red con elevadas prestaciones como, por ejemplo, una red WLAN o DVB. Así pues, el desarrollo de las redes heterogéneas promete ventajas para los usuarios, los proveedores de servicios y los proveedores de red. La implantación del concepto de red heterogénea llevará asociado un aumento en la capacidad de tráfico soportable, un incremento de los niveles de calidad de servicio ofrecidos a los usuarios y una reducción del coste económico asociado a la provisión de servicios. En las referencias [70–72] se muestran algunos estudios sencillos que ilustran estas ventajas.

Queda claro, por tanto, que el principal desafío para las redes de comunicaciones móviles del futuro consiste en proporcionar un amplio conjunto de opciones en términos de servicios, terminales de usuario y redes de acceso capaces de ofrecer flexibilidad, modularidad y escalabilidad. Sin embargo, antes de que ejemplos como el anterior puedan materializarse en el ámbito práctico, primero es necesario superar importantes dificultades que afectan tanto a los terminales de usuario como a los servicios y redes involucrados [73]. Estas dificultades derivan del hecho de que los diferentes estándares existentes no fueron concebidos para interoperar entre sí, sino que fueron diseñados con una arquitectura especializada de flexibilidad limitada. Para solventar este problema se

están proponiendo arquitecturas de interfuncionamiento (*interworking*) entre diversas tecnologías radio existentes [46–52, 74–76]. En este contexto, uno de los principales desafíos a los que se enfrentan los futuros sistemas heterogéneos de comunicaciones móviles es la habilidad de garantizar un adecuado nivel de calidad de servicio para el usuario mediante el empleo de diferentes tecnologías de acceso radio [77]. Esta nueva situación conlleva la necesidad de un interfuncionamiento adecuado entre dichas tecnologías para proporcionar una movilidad transparente y una buena calidad de servicio extremo a extremo [78–85]. La máxima calidad extremo a extremo que puede ofrecer cualquier red de comunicaciones viene limitada por el cuello de botella de dicha red, que en el caso de los sistemas de comunicaciones móviles es la interfaz radio. Por ello, uno de los elementos clave que debe considerar el proveedor de red es la gestión coordinada de los recursos radio (*Common Radio Resource Management*, CRRM) [86] de las distintas tecnologías como si de un único sistema se tratase, con el fin de lograr un aprovechamiento óptimo de los escasos recursos radioeléctricos y, en definitiva, una utilización más eficiente del sistema en su conjunto. La investigación de técnicas CRRM en redes heterogéneas es un área considerada en este momento como uno de los aspectos clave en el desarrollo hacia la 4G y figura entre las prioridades de algunos de los proyectos y redes de excelencia recientemente llevados a cabo dentro del VI Programa Marco de la Unión Europea.

El concepto de CRRM es relativamente amplio y abarca técnicas de diversa índole [86]. Una de ellas, por ejemplo, es la que recibe el nombre de asignación dinámica de espectro (*Dynamic Spectrum Allocation*, DSA) [87–95]. La motivación de esta técnica surge de la ineficiente utilización del espectro por parte de los sistemas actuales, en la medida en que éste es asignado mediante licencias estáticas que reservan cada banda de frecuencias a un estándar radio determinado, y estas bandas se dividen en fragmentos más pequeños que son asignados de forma estática a diversos operadores de una misma tecnología o servicio. Estos bloques de espectro tienen un tamaño constante y sólo son empleados por el operador que posee la licencia, hasta que ésta expira al cabo de varios años. Este procedimiento supone una forma sencilla de controlar la interferencia entre diversas redes radio y simplifica notablemente el *hardware* asociado, que se diseña para funcionar a una frecuencia fija. Sin embargo, supone un aprovechamiento ineficiente de los recursos radio, tan escasos como costosos. Frente a este esquema tradicional, las técnicas DSA proponen gestionar de forma conjunta el espectro empleado por diversas redes de acceso radio de una determinada red heterogénea de forma que éstas puedan compartirlo entre sí. La idea consiste en asignar bandas de frecuencia cuyo tamaño varía de forma dinámica para adaptarse a variaciones tanto temporales como espaciales experimentadas en la demanda, de forma que una misma frecuencia pueda ser utilizada por redes diferentes dependiendo del instante temporal y de la localización geográfica. Según el esquema DSA, cuando una determinada red experimenta un aumento de tráfico en una determinada zona o en una determinada franja horaria, puede hacer uso de parte del espectro de otra red con menor demanda de tráfico, con lo cual se mejora la eficiencia espectral del sistema en su conjunto. Esta técnica, si bien resulta interesante y bastante prometedora, presenta serias implicaciones no sólo en el ámbito tecnológico sino también en el ámbito regulatorio. Este último, desafortunadamente, avanza de forma considerablemente más lenta que el primero.

Otro aspecto importante dentro del ámbito de las técnicas CRRM es el traspaso de una comunicación en curso entre dos tecnologías radio diferentes, proceso que recibe el nombre de *traspaso o handover vertical* para diferenciarlo del traspaso entre celdas vecinas de una misma tecnología celular, denominado *traspaso o handover horizontal*. El interés de este tipo de técnicas se debe a que se prevé que el despliegue de cada tecnología de acceso radio se centre en un entorno espacial distinto. Las diferentes tecnologías, por lo tanto, se utilizarán de forma complementaria de modo que unas permitan suplir la carencia de cobertura de otras en determinadas zonas, mejorando así la cobertura global del sistema. Por ejemplo, mientras las redes WLAN (que soportan una baja movilidad de los usuarios pero ofrecen un elevado ancho de banda) se centrarían en *hot spots* [44] como aeropuertos, estaciones o grandes centros comerciales, las redes UMTS (con menores prestaciones pero mayor cobertura) serían desplegadas en zonas urbanas con alta densidad de tráfico y las redes GPRS (ampliamente implantadas pero con prestaciones más modestas) se centrarían en zonas suburbanas y rurales con menor densidad de tráfico (ver figura 1.3). En este contexto, el estudio de algoritmos de traspaso o *handover* vertical aborda principalmente la gestión de la movilidad y la calidad de servicio a medida que un usuario en movimiento realiza un traspaso (*handover*) entre diferentes sistemas de acceso radio debido a la falta de cobertura [96–98].

Sin embargo, este escenario facilita otro tipo de situaciones de mayor interés en las que se propone tener por objeto el gestionar de manera óptima la disponibilidad de diversas tecnologías de acceso radio en una localización geográfica determinada, más allá de la gestión de la movilidad intersistema. En el estudio de este tipo de situaciones, el traspaso de un usuario entre dos tecnologías radio diferentes no es necesariamente un proceso que ha de estar determinado únicamente por la movilidad del usuario, sino que puede decidirse también en función de otros muchos factores como, por ejemplo, la disponibilidad del enlace, la calidad del enlace, el estado de las redes, los requisitos de calidad de servicio, las preferencias de los usuarios, el coste económico o el impacto de la nueva comunicación. Para entender la necesidad y la importancia de gestionar adecuadamente este tipo de situaciones supongamos, a modo de ejemplo, un usuario cuyo terminal puede funcionar indistintamente bajo una red celular o bajo una red WLAN. Inicialmente el terminal de dicho usuario se halla conectado a una red celular, pero al entrar a un aeropuerto detecta la presencia de una señal más fuerte proveniente de una red WLAN, cuyas prestaciones en términos de régimen binario además son mayores que las de la red celular. Ante esta situación, el comportamiento lógico que cabría esperar por parte del terminal del usuario sería el de desconectarse de la red celular y conectarse a la red WLAN del aeropuerto mediante un traspaso o *handover* vertical. Si todos los terminales de usuario reaccionan de la misma manera, el resultado más probable es que la red WLAN acabe fuertemente cargada, incluso saturada, mientras que la red celular queda infrautilizada dentro del aeropuerto. Esto no sólo supone un uso claramente ineficiente de los recursos radio disponibles, sino también la existencia de una red cuyo coste de operación y mantenimiento permanece constante pero cuyos beneficios disminuyen. Queda claro, así pues, la necesidad e importancia del estudio de algoritmos que hagan un uso más racional de las diferentes tecnologías de acceso radio disponibles en una determinada localización, distribuyendo el tráfico entre ellas de forma inteligente para así lograr un aprovechamiento óptimo del espectro disponible.

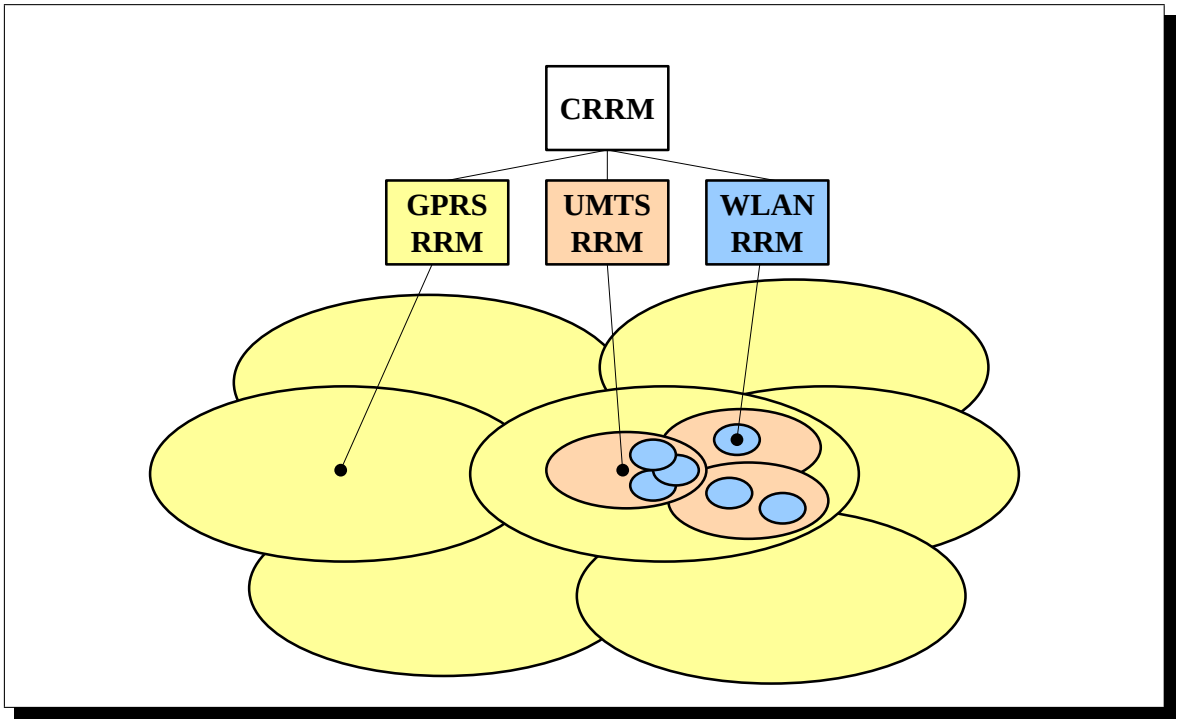
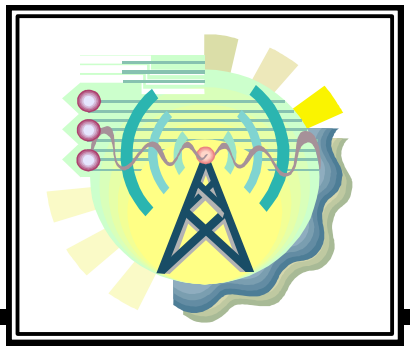


Figura 1.3: Ejemplo de cobertura jerárquica en una red heterogénea.

El interés del presente trabajo recae precisamente sobre el estudio de este tipo de técnicas, que en la literatura científico-técnica ha recibido diferentes denominaciones como distribución de tráfico, control de tráfico, selección de acceso radio (*Radio Access*, RA), selección de red de acceso radio (*Radio Access Network*, RAN) o selección de tecnología de acceso radio (*Radio Access Technology*, RAT). Si bien existen varios trabajos previos en la materia, este campo de investigación es relativamente reciente y todavía requiere mucho más esfuerzo del que se le ha dedicado hasta ahora. En este contexto, el presente trabajo propone y evalúa nuevos algoritmos CRRM diseñados para la distribución eficiente de tráfico entre las diferentes alternativas de acceso radio disponibles en un entorno heterogéneo de comunicaciones móviles, dando respuesta al problema de asignar los usuarios del sistema a aquella red de acceso radio que resulte óptima. El objetivo último de este trabajo es el de proveer soluciones que permitan satisfacer los diversos requisitos de calidad de servicio de una red heterogénea al mismo tiempo que se garantiza un aprovechamiento óptimo de los recursos radio disponibles.

El resto del trabajo se organiza según se detalla a continuación. En primer lugar, en el capítulo 2 se realiza una revisión del trabajo previo realizado por otros autores así como del llevado a cabo en el seno de diversos proyectos de investigación y estandarización de ámbito europeo. A continuación, el capítulo 3 presenta y describe las soluciones propuestas en este trabajo. El rendimiento de estas propuestas se evalúa mediante simulaciones a nivel de sistema llevadas a cabo con una avanzada plataforma de simulación, cuya descripción tiene lugar en el capítulo 4. Los resultados obtenidos se presentan y analizan en el capítulo 5. Finalmente, el capítulo 6 expone las principales conclusiones derivadas de la investigación llevada a cabo en este trabajo.



CAPÍTULO 2

Trabajo Previo Relacionado

El objetivo de este capítulo es el de realizar una revisión del trabajo previo de otros autores que guarda relación, de manera más o menos directa, con el problema de distribuir el tráfico de una red heterogénea entre sus diferentes alternativas de acceso radio, alternativas que en lo sucesivo denominaremos, indistintamente, redes de acceso radio (*Radio Access Networks*, RANs) o tecnologías de acceso radio (*Radio Access Technologies*, RATs). La solución al problema de la distribución de tráfico en una red heterogénea requiere, por una parte, la definición de algoritmos que describan de manera clara un procedimiento para decidir cómo asignar cada usuario a una de las RANs disponibles y, por otra parte, la especificación de una arquitectura lógica de gestión (arquitectura CRRM) que dé soporte a las funcionalidades requeridas por dichos algoritmos y facilite la obtención de la información requerida por el criterio de decisión aplicado. La especificación de dicha arquitectura queda fuera de los objetivos de este trabajo y, por lo tanto, en el presente capítulo se omitirá una revisión del trabajo realizado al respecto. El lector interesado puede encontrar algunos ejemplos de arquitecturas CRRM en las referencias [69, 70, 99–102]. El interés de este trabajo, tal y como se señaló en el capítulo 1, se centra sobre la definición de los algoritmos que deben decidir la distribución de tráfico sobre las RANs que integran una red heterogénea. Consecuentemente, los trabajos citados en este capítulo tienen en común la definición de algún principio o criterio para decidir dicha distribución de tráfico. El objetivo de este capítulo no es el de llevar a cabo una revisión exhaustiva del trabajo realizado en este campo ni el de analizar en detalle los criterios propuestos en otros trabajos, sino más bien el de introducir la problemática asociada a los objetivos de este trabajo y presentar algunos criterios y soluciones previamente propuestos por otros autores.

2.1. Revisión de la literatura

Aunque haya sido considerado como un aspecto de gran importancia para la gestión de recursos radio en redes heterogéneas B3G/4G, la idea de distribuir el tráfico entre diferentes RANs no es exclusiva del concepto de red heterogénea. Históricamente, la posibilidad de reducir los costes de infraestructura y mejorar la flexibilidad de la red mediante la distribución de tráfico fue demostrada por primera vez en una red dual AMPS/DAMPS, en la cual los terminales DAMPS capaces de operar en modo dual podían conectarse a la red AMPS y los terminales AMPS, anteriores a los terminales DAMPS, podían conectarse a las estaciones base DAMPS capaces de operar en modo dual [103–106]. La idea de distribución de tráfico también fue propuesta para los sistemas estadounidenses de primera y segunda generación AMPS e IS-95 respectivamente, como medio para mejorar la flexibilidad del sistema y reducir los costes de infraestructura [107]. En estos trabajos se supone la existencia de un único servicio (el servicio de voz) y de una única banda de frecuencias compartida. Las redes heterogéneas B3G/4G pueden operar en distintas bandas de frecuencia y, aunque pueden proporcionar servicios de voz y datos, su foco de interés se centra sobre estos últimos.

A la hora de proponer un criterio de distribución de tráfico en una red heterogénea hay que tener en cuenta que las RANs que integran dicha red poseen, en general, unas características particulares y específicas que, en cierta forma, las hacen más o menos indicadas para un determinado tipo de usuario o servicio. Resulta interesante citar al respecto la referencia [108]. En este trabajo se describen las principales características de algunas de las RANs más extendidas y que están llamadas a formar parte de las futuras redes heterogéneas como son GERAN (*GSM/EDGE Radio Access Network*), UTRAN (*UMTS Terrestrial Radio Access Network*), WLAN y Bluetooth. También realiza un estudio de las prestaciones de rango de cobertura, capacidad y retardo de dichas RANs. Ignorando aspectos comerciales como coste, facturación, *roaming* o seguridad y teniendo en cuenta únicamente estos tres parámetros tecnológicos (rango, capacidad y retardo), el artículo realiza algunas indicaciones acerca del conjunto de RANs que pueden resultar más indicadas en función de las características del tipo de servicio considerado. Si bien resultan interesantes las observaciones realizadas en [108], el artículo no propone formalmente ningún principio o criterio concreto.

Comenzaremos la revisión de este capítulo presentando varios trabajos que proponen algoritmos de distribución de tráfico sencillos. En principio, el tipo de algoritmo más sencillo que puede aplicarse es el consistente en asignar, de manera sistemática, un determinado grupo de usuarios (establecido en función del tipo de servicio o según otro criterio) siempre a la misma RAT. Los usuarios de ese grupo siempre intentarán conectarse preferiblemente a la RAT preestablecida y, en caso de fracasar el intento, probarán consecutivamente con el resto de RATs disponibles siguiendo un orden establecido según cierto criterio. Un ejemplo de este tipo de algoritmos es el trabajo presentado en la referencia [109]. En este artículo se proponen diversos algoritmos sencillos aplicables a la selección inicial de RAT, es decir, aplicables a la selección de la RAT a través de la cual debe transmitir un nuevo usuario en el momento en el que inicia una nueva sesión de comunicación. En un escenario formado por una red GERAN

y una red UTRAN, este trabajo considera la selección inicial de RAT para usuarios de dos tipos de servicio: voz y navegación web. Inicialmente se proponen tres principios de asignación básicos. El primero asigna, de manera sistemática, el tráfico de voz a GERAN y el tráfico web a UTRAN. El segundo de ellos realiza la asignación justamente al revés, es decir, el tráfico de voz se encamina a través de UTRAN mientras que el tráfico web se encamina a través de GERAN. Un tercer principio propuesto consiste en asignar los usuarios en entornos interiores (*indoor*) a GERAN y los usuarios en entornos exteriores (*outdoor*) a UTRAN. Este último principio de asignación se propone dada la degradación de la capacidad que experimenta un sistema WCDMA, como es el caso de UTRAN, ante la presencia de usuarios en entornos interiores, según se señala en [110]. Estos tres principios básicos muestran todos ellos el inconveniente de que un determinado tipo de usuario siempre intentará conectarse a la misma RAT, tanto si ésta dispone de capacidad suficiente para aceptarlo como si no. Consecuentemente, estos principios básicos implican la existencia de cierta probabilidad de bloqueo de los usuarios entrantes, aun cuando existe capacidad suficiente para darles servicio en alguna de las RATs alternativas. Por ello, en este trabajo se proponen otros principios alternativos consistentes en combinar, de manera consecutiva, dos de los tres principios básicos iniciales. A modo ilustrativo, se describirá el principio de asignación empleado en [111], el cual consiste en aplicar en primer lugar el primero de los principios básicos (voz a GERAN y web a UTRAN) y, si éste fracasa, se aplica en segundo lugar el segundo de los principios básicos (voz a UTRAN y web a GERAN). Según este principio de asignación compuesto, un usuario de voz intenta la conexión a GERAN y, en caso de ser rechazado, intenta la conexión a UTRAN. De manera similar, un usuario web intenta primero la conexión a UTRAN y, si resulta rechazado, entonces intenta la conexión a GERAN. Si un usuario es rechazado por las dos RATs disponibles, entonces queda bloqueado. El funcionamiento del algoritmo se ilustra en la figura 2.1, donde se ha añadido la posibilidad de que los terminales del sistema sean monomodo, es decir, capaces de operar en una única RAT (GERAN en este caso). La idea general que subyace bajo este planteamiento puede ser fácilmente extendida y aplicada a un número arbitrario de RATs y de servicios.

Son varias las críticas que pueden realizarse al algoritmo de la figura 2.1. Una de ellas es que la asignación de los usuarios se realiza sin considerar variaciones temporales de las condiciones de operación como, por ejemplo, la carga existente en cada RAT. Un usuario de voz intentará conectarse primero a GERAN, aunque la red GERAN se halle fuertemente cargada y la red UTRAN disponga de capacidad libre para nuevos usuarios. Esto no sólo supone una degradación innecesaria de la calidad del resto de usuarios de GERAN, sino también en algunas ocasiones un cierto tráfico de señalización, asociado al proceso de solicitud y rechazo de establecimiento en GERAN, que podría haberse evitado intentando directamente una conexión a UTRAN. Otro inconveniente de esta propuesta, que hace cuestionable su eficiencia, es que puede darse la situación de que un terminal monomodo que sólo pueda operar en GERAN quede bloqueado si dicha RAN está saturada, situación que puede evitarse traspasando un usuario multimodo de GERAN a UTRAN mediante un *handover* vertical con la finalidad de liberar los recursos requeridos por el nuevo terminal monomodo, el cual podría entonces ser admitido en GERAN una vez realizado el traspaso.

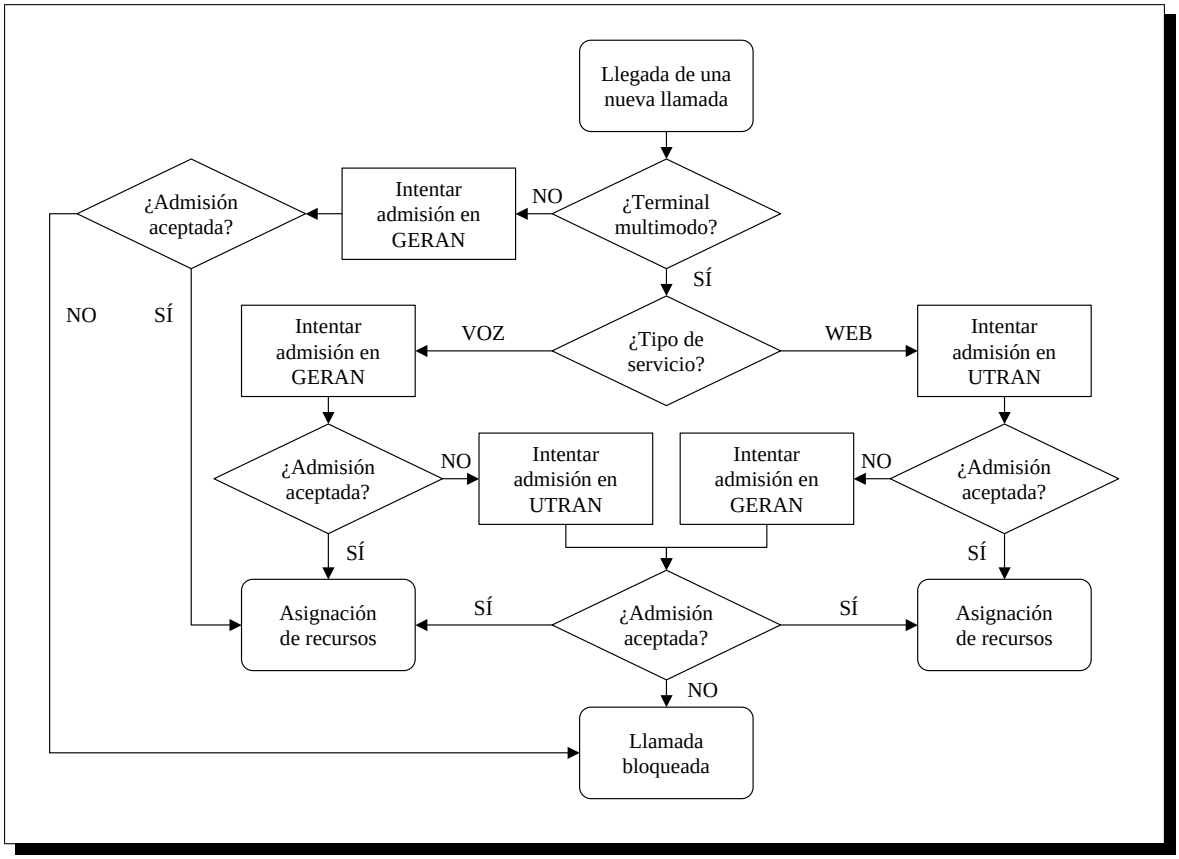


Figura 2.1: Ejemplo de algoritmo basado en asignaciones preestablecidas [109, 111].

Esta es precisamente la idea en la cual se basa la propuesta realizada en [112, 113]. En dichos artículos se proponen varios criterios de distribución de tráfico también basados en el principio de intentar inicialmente una asignación preestablecida, pero que supera los inconvenientes del algoritmo de la figura 2.1. De manera similar, se propone intentar una asignación inicial de los usuarios de voz a GSM y de los usuarios de datos a UMTS. En el caso de que ante la llegada de una nueva llamada de voz o de datos no exista capacidad suficiente en la red GSM o UMTS respectivamente, entonces no se intenta una conexión a la RAT alternativa, sino que se propone la aplicación de un determinado criterio de traspaso de usuarios entre ambos sistemas. El objetivo de realizar este traspaso de usuarios es el de habilitar la capacidad necesaria en la RAT saturada para la nueva llamada entrante, intentando así maximizar la capacidad global del sistema. Dicho criterio se caracteriza por una política de traspaso y una política de retorno de los usuarios. Se consideran dos posibles políticas de traspaso, una *simple* en la cual se intenta traspasar la última llamada que ha llegado a la RAT saturada, pero sólo si dicho terminal soporta la operación multimodo (equivalente a intentar establecer una conexión con la RAT alternativa), y otra *compleja* en la cual se intenta traspasar cualquier llamada establecida por un terminal multimodo en caso de que la nueva llamada sea solicitada por un terminal monomodo. Adicionalmente, se consideran dos posibles políticas de retorno, una de *retorno* en la cual los usuarios traspasados desde una RAT por falta de capacidad regresan a dicha RAT una vez que vuelve a existir

cierta capacidad disponible, y otra de *no retorno* en la cual los usuarios que son traspasados ya no regresan a la RAT de origen. Este planteamiento da lugar a cuatro posibles criterios: simple sin retorno, simple con retorno, complejo sin retorno y complejo con retorno. De estos cuatro, el último de ellos ofrece las mejores prestaciones. Observar que esta propuesta es una generalización de la anterior, donde el algoritmo de la figura 2.1 corresponde al criterio simple sin retorno. En [114] se introduce una política de retorno adicional, cuyo objetivo consiste en reducir la frecuencia de los *handovers* verticales y, por tanto, la carga de señalización asociada al algoritmo de distribución de tráfico. Dicha política de retorno consiste en que los terminales traspasados vuelven a la RAT de origen, no cuando vuelve a existir capacidad disponible en ella, sino cuando la RAT de destino se satura. Los criterios anteriores suponen un método sencillo para conseguir una mayor capacidad en la red y una rápida adaptabilidad a las fluctuaciones temporales del tráfico. Según muestran los resultados obtenidos en [113], con la aplicación de estos algoritmos, el aumento de capacidad asociado a una única celda no resulta significativo, si bien sí puede serlo si los algoritmos se implementan en toda la red. Otros trabajos relacionados con estas propuestas son las referencias [115, 116], donde se realiza un estudio ampliado bajo diversas condiciones de operación, o [117, 118], donde se estudian y clasifican metódicamente los diferentes algoritmos que pueden derivarse de las diferentes posibles combinaciones de criterios anteriormente citados. En [119] se discute la posibilidad de obtener ganancias adicionales en la capacidad del sistema combinando los algoritmos anteriores, basados en el traspaso de llamadas, con la asignación dinámica de espectro y la degradación intencionada de la calidad de llamadas ya establecidas para poder aceptar nuevas llamadas.

En [120] se presenta otra propuesta de algoritmo para la distribución de tráfico en redes heterogéneas donde las decisiones se toman, al igual que en el caso anterior, en función de la carga de cada una de las RANs del sistema. Según se discute en dicho artículo, resulta deseable obtener una distribución uniforme del tráfico entre las diversas RANs que integran el sistema para obtener así una mayor capacidad de usuarios y una menor probabilidad de traspasos intersistema (*handovers* verticales), lo cual implica una reducción de la carga de señalización asociada. Una distribución uniforme del tráfico no implica igualar la carga soportada por cada una de las RANs del sistema, sino evitar que una de ellas se encuentre altamente cargada mientras existen otras que soportan una carga de tráfico reducida. En el citado trabajo se proponen diferentes algoritmos en función de si el tipo de servicio es en tiempo real o no lo es.

Para servicios en tiempo no real, el algoritmo propuesto en [120] se basa en el hecho de que es habitual la utilización de un *buffer* de transmisión en cada celda del sistema en el cual los paquetes de datos de los usuarios se colocan antes de ser transmitidos y son servidos según cierta política de servicio o de *scheduling*. En este artículo se considera que los paquetes de los usuarios se transmiten según el esquema *First-In First-Out* (FIFO), en ocasiones también referido como *First-Come First-Served* (FCFS), es decir, en el mismo orden en que llegan al *buffer* de transmisión. Cuando el tiempo medio de espera de los paquetes en el *buffer* supera cierto valor máximo, entonces se busca la RAN para la cual el valor de dicho parámetro es el menor de todos. Si la diferencia entre los valores para la RAN actual y para la RAN con menor retardo medio supera cierto

umbral, entonces se sustrae el último paquete de datos insertado en el *buffer* de la RAN actual y se coloca en el de la RAN con el menor retardo medio, lo cual significa mover un usuario de una RAN a otra. El umbral de retardo máximo de *buffer* se establece para evitar que la frecuencia de los *handovers* verticales sea elevada, mientras que el umbral de diferencia de retardos entre RANs se establece para prevenir *handovers* verticales (con la señalización que ello conlleva) en los cuales las mejoras obtenidas son poco significativas. Este algoritmo resulta aplicable a la decisión de *handovers* verticales una vez que los usuarios ya han sido asignados a una determinada RAN. Para usuarios que inician una llamada, se propone escoger aquella RAN para la cual el retardo medio del *buffer* de transmisión es menor, es decir, aquella RAN con la menor carga de tráfico.

Para servicios en tiempo real, el algoritmo propuesto en [120] es completamente diferente al propuesto para servicios en tiempo no real, aunque el objetivo perseguido es el mismo. En el caso de servicios en tiempo real, los usuarios permanecen asignados a su RAN actual mientras la carga de tráfico permanezca por debajo de cierto umbral. Si la carga de la RAN supera el umbral establecido pero no alcanza el 100 %, entonces se intenta realizar un *handover* vertical, de aquellos usuarios que provocan un exceso de carga, hacia la RAN menos cargada. Si para dicha RAN la carga no supera el umbral, entonces el *handover* se realiza, pero en caso contrario el *handover* no tiene lugar y las llamadas de estos usuarios se siguen manteniendo en la RAN de origen. Cuando la carga de la RAN de origen alcanza el 100 %, entonces se intenta traspasar a aquellos usuarios que no tienen cabida hacia la RAN de destino. Si ésta también se halla al máximo de sus posibilidades, entonces los usuarios quedan bloqueados. En caso contrario son aceptados en la RAN de destino, independientemente de que la carga de ésta supere o no el umbral. Este procedimiento determina la decisión de los *handovers* verticales. Para la selección inicial de RAN se propone elegir aquella menos cargada.

En la referencia [121], los autores proponen una mejora al planteamiento anterior consistente en ajustar dinámicamente los umbrales que determinan el disparo del proceso de *handover* vertical. Proponen ajustar estos umbrales en función de la carga de las RANs alternativas de destino para evitar *handovers* innecesarios cuando las RANs del sistema están altamente cargadas. Por ejemplo, si en una RAN determinada se establece el umbral al 80 % de la carga máxima soportable y se supera dicho umbral, pero las demás RANs soportan un nivel de carga igual al 90 % del máximo, no tiene sentido intentar realizar un *handover* vertical hacia una RAN más cargada. Ajustando el valor de estos umbrales para cada RAN en función de la carga soportada por las RANs alternativas se consigue reducir el número de intentos de *handover* innecesarios.

Los trabajos presentados hasta el momento proponen criterios basados en asignaciones preestablecidas de determinados tipos de usuario a cada RAN o cuyas decisiones se basan en la carga actual de cada RAN. No obstante, en ninguna de las propuestas anteriores la asignación de usuarios a una de las RANs disponibles se realiza teniendo en cuenta el nivel de calidad de servicio requerido por el usuario ni las prestaciones o calidad de servicio que puede ofrecer cada RAN. En contraste con las propuestas anteriores, en [122] se proponen diversos criterios de distribución de tráfico. Algunos de ellos también basan sus decisiones en la carga actual de cada RAN, aunque otros basan su decisión en el nivel de calidad de servicio que puede esperarse de cada RAN.

En [122] se proponen en total cinco algoritmos pensados para asignar cada usuario del sistema a una de las RANs disponibles. En este caso el escenario considerado está formado por una red celular WCDMA y una red WLAN IEEE 802.11. Los cinco algoritmos propuestos se clasifican en dos grupos. Los algoritmos que pertenecen al primero de estos grupos seleccionan la red WLAN si la relación señal a ruido (*Signal-to-Noise Ratio*, SNR) de dicha red supera cierto valor mínimo. En caso contrario seleccionan la red celular. Dentro de este grupo se proponen tres algoritmos diferentes. El primero de ellos establece un valor mínimo de SNR igual a 0 dB, es decir, selecciona la red WLAN siempre que existe cobertura de dicha red. El segundo establece el valor mínimo de SNR de tal forma que se obtenga un equilibrio entre la carga relativa de cada RAN, es decir, que el valor del cociente entre el número de usuarios y la capacidad sea el mismo para todas las RANs. El tercero consiste en establecer el valor mínimo de SNR en función de la carga de la red WLAN. Se asigna un valor inicial de 0 dB cuando la carga de la red WLAN es baja y se va incrementando dicho valor a medida que el tráfico de la red WLAN aumenta (y viceversa), con la finalidad de controlar la carga de la red WLAN manteniéndola por debajo de su capacidad máxima. En este primer grupo de algoritmos, el objetivo es en cierta forma el de controlar la carga de cada RAN, ajustando para ello el valor mínimo de SNR requerido de la red WLAN. El segundo grupo de algoritmos sugiere basar las decisiones en el nivel de calidad de servicio que se puede esperar de cada RAN, considerando la velocidad de transmisión de datos como parámetro de calidad. Los algoritmos de este grupo, por tanto, seleccionan aquella RAN que ofrece la mayor velocidad de transmisión. El primer algoritmo de este grupo estima la velocidad de transmisión de cada RAN a partir de la tasa binaria experimentada en el enlace radio, obtenida únicamente a partir del nivel de calidad del enlace radio. El segundo algoritmo de este grupo estima la velocidad de transmisión a partir de la tasa equivalente en conmutación de circuitos (*Circuit Switched Equivalent*, CSE), la cual considera adicionalmente la carga del sistema. Los resultados de este estudio muestran que, en general, el criterio más sencillo de todos, el consistente en seleccionar la red WLAN siempre que ofrezca cobertura, ofrece en general resultados satisfactorios. Sin embargo, en escenarios en los cuales existen tecnologías celulares cuyas prestaciones en términos de velocidad de transmisión son del mismo orden que las de la red WLAN, o en escenarios con *hot spots* de elevada carga de tráfico, se obtienen mejores resultados mediante la aplicación de criterios más sofisticados en los cuales se tiene en cuenta la carga del sistema. Estos resultados recuerdan las afirmaciones realizadas en el ejemplo del capítulo 1 a raíz del ejemplo del aeropuerto.

En [123] también se proponen varios algoritmos cuyo criterio sugiere que cada usuario seleccione aquella RAN que ofrece la máxima tasa binaria de transmisión. En el citado trabajo se distingue entre dos tipos de principios de selección. Uno de ellos está basado en la perspectiva del terminal de usuario, siendo cada terminal quien selecciona la RAN óptima, considerándose como óptima aquella que proporciona la máxima tasa binaria de transmisión esperada en función de la calidad del canal. El otro principio está basado en la perspectiva de la red, donde cada RAN es quien elige al usuario a servir en cada momento, seleccionando aquél que maximiza el cociente entre la tasa de transmisión esperada y la tasa de transmisión requerida por el usuario. Tal y como se discute en dicho trabajo, la estrategia basada en la perspectiva de la red se beneficia de

la diversidad multiusuario (al poder elegir entre cualquiera de los usuarios del sistema) y de la diversidad multiradio (al poder estar conectado un usuario de forma simultánea a múltiples RANs, siempre y cuando éstas lo hayan seleccionado como usuario preferido). En cambio, la estrategia basada en la perspectiva del terminal de usuario sólo puede explotar la diversidad multiradio (al poder un usuario elegir entre varias RANs). Los resultados muestran que, para ambos principios, es posible obtener ciertas ganancias gracias a la diversidad multiusuario y multiradio respecto al caso en el que no se aplica ningún criterio de selección de acceso radio y, por tanto, cada usuario siempre se conecta a la misma RAN. La estrategia más eficiente es la basada en la perspectiva de la red, en la cual se explota la diversidad tanto multiusuario como multiradio.

Continuaremos la revisión de este capítulo introduciendo ahora un enfoque diferente a los anteriores para elaborar algoritmos de distribución de tráfico en redes heterogéneas. Este enfoque consiste en definir una cierta función denominada *función objetivo* o *función de costes*. Tal y como se verá en breve, existen diferentes formas de utilizar dicha función para decidir una distribución de tráfico determinada. Una de ellas, por ejemplo, consiste en definir una función que asigne un cierto valor numérico a cada una de las RANs disponibles en el sistema. La decisión puede realizarse entonces escogiendo la RAN que ofrece el valor más elevado para dicha función, en cuyo caso suele denominarse función objetivo ya que el objetivo es el de maximizar el valor de dicha función, o escogiendo la RAN que ofrece el valor más reducido, en cuyo caso suele denominarse función de costes ya que el objetivo es el de minimizar el valor de dicha función. En este tipo de planteamientos, la definición del algoritmo requiere, por una parte, definir la expresión de la función y el procedimiento seguido para calcular el valor de cada uno de sus parámetros y, por otra parte, decidir en qué instante durante el funcionamiento del sistema se evalúa dicha función para tomar una decisión concreta.

En primer lugar, una expresión sencilla y común a varios trabajos que se basan en este tipo de planteamientos es la siguiente

$$f(c,p) = \sum_{i=1}^{N} c_i \cdot p_i \tag{2.1}$$

en cuyo cómputo interviene un número arbitrario N de parámetros $p = \{p_1, p_2, \ldots, p_N\}$ cuya suma ponderada por los coeficientes $c = \{c_1, c_2, \ldots, c_N\}$ proporciona el valor de la función objetivo $f(c,p)$. Cada uno de estos N parámetros suele calcularse a partir de diversas variables y puede tomar valores en un cierto rango predefinido.

El trabajo presentado en la referencia [124] es un ejemplo basado en una función objetivo de la forma indicada en la expresión 2.1. Los parámetros que intervienen en el cálculo de dicha función representan el nivel de calidad de servicio, el coste económico del servicio, el operador de red preferido y el tipo de RAN preferida. La idea que propone el citado trabajo consiste en que sea el usuario el que especifique, mediante una interfaz gráfica en su terminal móvil, el coeficiente de ponderación de cada uno de estos parámetros. De esta forma el usuario indica sus preferencias mediante el valor de los coeficientes de ponderación, cuantificando la importancia relativa que asigna a cada

uno de los parámetros anteriores a la hora de seleccionar una determinada RAN. Cada terminal seleccionará la RAN idónea evaluando la función objetivo para cada una de las RANs candidatas y escogiendo aquella que ofrece el máximo valor. En el artículo se proponen expresiones para calcular el valor numérico asociado a los parámetros de calidad de servicio y coste. El valor numérico del parámetro de operador de red se establece asignándole un cierto porcentaje del valor absoluto de los parámetros calidad y coste cuando el operador de la RAN candidata es el previamente especificado por el usuario como preferido, o cero en caso contrario. El valor numérico del parámetro de tipo de RAN se establece de forma similar. El otro aspecto a definir por el algoritmo, tal y como se señaló anteriormente, es cuándo debe evaluarse la función objetivo para seleccionar la RAN más adecuada. En el citado trabajo se propone evaluar la función objetivo en cuatro circunstancias diferentes: cuando se solicita un nuevo tipo de servicio, cuando un usuario modifica su perfil de preferencias, cuando se detecta una variación significativa en las condiciones de propagación radio (disminución del nivel de señal recibido) y cuando se detecta la disponibilidad de un nuevo punto de acceso radio.

Otro ejemplo de propuesta basada en la evaluación de una función objetivo de la forma indicada en la expresión 2.1 es la realizada en [125]. En dicho trabajo se propone aplicar ideas y planteamientos previamente utilizados dentro del área de sistemas de cómputo distribuido. Un sistema de cómputo distribuido es un sistema en el cual existen varios elementos individuales de cómputo (procesadores), entre los cuales se realiza un reparto de los trabajos o tareas a realizar. Puede observarse cierta similitud entre el problema de distribuir un determinado número de tareas a realizar entre diferentes unidades de cómputo y el problema de distribuir un determinado número de terminales móviles entre diferentes RANs dentro de una red heterogénea. No obstante, existen importantes diferencias entre ambos problemas [125] que impiden aplicar directamente los algoritmos empleados en sistemas de cómputo distribuido al problema de la distribución de tráfico en redes heterogéneas. Los autores proponen, por lo tanto, un algoritmo específico para su aplicación en redes heterogéneas pero basado en los principios de los *modelos de fuerzas* (*force models*), previamente utilizados en el diseño de algoritmos en sistemas de cómputo distribuido. En última instancia, el algoritmo propuesto se reduce a evaluar, para cada RAN, una función objetivo de la forma indicada en la expresión 2.1, en la cual intervienen diversos parámetros o *fuerzas* (usando la terminología del modelo empleado). Cada usuario selecciona la RAN que ofrece el valor más elevado para dicha función. En su cálculo, los parámetros que intervienen son cuatro. El primero de ellos representa una estimación de la capacidad libre en cada una de las RANs alternativas. El segundo representa la diferencia entre la calidad de servicio esperada en cada RAN alternativa y la experimentada en la RAN actual, calidad expresada mediante la tasa de errores o la tasa binaria de transmisión de datos. El tercer parámetro representa el tiempo transcurrido entre el último *handover* vertical realizado y el instante de evaluación de la función, parámetro incluido con el objetivo de limitar la frecuencia máxima de ejecución de *handovers* verticales así como evitar el indeseable efecto *ping-pong*, que tiene lugar cuando un terminal realiza traspasos muy frecuentes entre dos RANs diferentes. El cuarto y último parámetro considerado representa el coste asociado al traspaso intersistema, es decir, la carga de señalización derivada de la realización de un *handover* vertical. Las expresiones y procedimientos

empleados para asignar valores a estos parámetros se describen en el propio artículo. Los autores proponen medir de forma periódica la carga de cada RAN. En el momento en que se supera un cierto umbral de carga, entonces se realizan las medidas necesarias para obtener los valores asociados a cada parámetro que interviene en la función. Una vez disponible la información necesaria para evaluar la función, se procede a su cálculo y, en caso de ser necesario, se procede a realizar los *handovers* verticales pertinentes. Este proceso de traspaso de usuarios continúa hasta que la carga de tráfico de la RAN considerada queda por debajo de un determinado valor, el cual se obtiene restando un cierto margen de histéresis al umbral de carga máxima. En la definición del algoritmo los autores también especifican una política de prioridades a la hora de realizar el traspaso de los usuarios, la cual consiste en atender primero a aquellos usuarios con un mayor consumo de recursos por bit (ocasionado por unas condiciones radio menos favorables). Los resultados mostrados indican un comportamiento bastante estable del algoritmo en situaciones de sobrecarga y una buena capacidad de respuesta ante cambios lentos y rápidos en la carga de tráfico soportada.

Aunque originalmente no se plantea como tal, el trabajo presentado en [126] puede, en cierta forma, englobarse dentro del conjunto de propuestas basadas en la evaluación de una función objetivo de la forma indicada en la expresión 2.1. En el citado trabajo se propone un algoritmo de selección de RAN basado en la aplicación conjunta de los principios de la lógica difusa y las redes neuronales. Por una parte, la lógica difusa ha demostrado ser eficaz en la explicación de la obtención de decisiones a partir de información imprecisa, como es el caso en una red heterogénea. En una red heterogénea, la selección de una determinada RAN se realiza en función de parámetros que proceden de RANs generalmente bastante diferentes y, por tanto, en función de parámetros que no son directamente comparables entre sí. Algunos ejemplos de este tipo de parámetros son, por ejemplo, el nivel de señal portadora de cada RAN, la carga de las celdas o la velocidad del terminal móvil [126]. Sin embargo, existen ciertos aspectos de la lógica difusa que resultan subjetivos y cuya determinación puede verse facilitada mediante la aplicación de redes neuronales, capaces de reconocer patrones mediante el empleo de técnicas de aprendizaje. En este aspecto, la lógica difusa y las redes neuronales resultan complementarias. De hecho, la aplicación de sistemas híbridos basados en lógica difusa y redes neuronales ha sido propuesta en varios campos como metodologías complementarias entre sí. En este artículo, se propone este tipo de sistemas como candidato para la selección de la RAN más adecuada para cada usuario dentro de una red heterogénea. Sin ánimo de entrar a describir esta propuesta en detalle, aquí simplemente se señalarán los aspectos cualitativos que caracterizan el proceso de decisión, señalando los parámetros utilizados para seleccionar una determinada RAN. La decisión de la RAN idónea se realiza en función de tres parámetros diferentes. El primero de ellos es un parámetro que establece, para cada una de las RANs disponibles, un valor numérico obtenido a partir de tres variables de entrada: el nivel de señal recibido, la disponibilidad de recursos y la velocidad del terminal móvil. Aplicando un sistema híbrido basado en lógica difusa y redes neuronales sobre estas tres variables, se obtiene el valor del primer parámetro. Este primer parámetro determina cuál de las RANs disponibles resulta idónea desde un punto de vista meramente tecnológico. El segundo de los parámetros utilizados representa el nivel de demanda en función del precio y de la

utilidad percibida por el usuario. El valor de este segundo parámetro se obtiene según la curva de Cobb-Douglas, bien conocida en Economía. El tercer parámetro representa las preferencias del operador y se establece de forma un tanto subjetiva en función de la política del operador y los acuerdos establecidos con otros operadores. Ponderando el valor de estos tres parámetros según se indica en la expresión 2.1, se obtiene un único valor numérico para cada una de las RANs disponibles, escogiéndose aquella que ofrece el valor más elevado. En este trabajo se describe el procedimiento seguido para obtener el valor numérico asociado a cada RAN, pero no se especifica cuándo o en qué circunstancias se calcula dicho valor para decidir la RAN a seleccionar.

En [56] también se realiza otra propuesta basada en la evaluación de una función objetivo o función de costes. Dicho trabajo no indica de forma explícita la expresión utilizada para la función evaluada, aunque habla de coeficientes de ponderación para los diversos parámetros que intervienen en su cálculo, lo cual hace suponer una expresión similar a 2.1. Tampoco se especifica cuándo se evalúa dicha función para tomar las decisiones pertinentes. En el citado trabajo se sugiere el empleo de diferentes parámetros en la función utilizada dependiendo de si la selección de RAN corresponde a una asignación inicial o a un *handover* vertical. Para la asignación inicial de RAN, la función propuesta incluye los valores asociados al tipo de terminal de usuario (teléfono móvil, PDA, ordenador portátil, ...), requisitos del tipo de tráfico a transmitir (velocidad mínima, retardo máximo, prioridad, seguridad requerida, ...), velocidad del terminal de usuario y preferencias del usuario (basadas en las experiencias previas del usuario, la especificaciones de subscripción, las conexiones existentes, ...). Para la decisión de los *handovers* verticales, la función empleada pondera los valores asociados a la velocidad de transmisión, la calidad del enlace radio expresada en términos de la tasa de paquetes erróneos, la movilidad del usuario y la carga de tráfico. A partir del valor obtenido mediante estas funciones se decide la RAN óptima.

Existen trabajos donde la función a evaluar se define como una función de costes que asigna un coste a cada una de las posibles alternativas, escogiéndose aquella para la cual el valor de dicho coste es el menor de todos. Este es el caso de la propuesta realizada en [127], donde se considera un escenario formado por GPRS y UMTS. Cada RAN calcula el coste asociado a cada uno de los servicios prestados en el sistema para los diferentes niveles de calidad de servicio con los que éstos se prestan. El valor de estos costes se difunde entre los terminales móviles del sistema. Cada terminal móvil de usuario recibe esta información de cada una de las RANs que le ofrecen cobertura y, teniendo en cuenta dicha información así como sus preferencias relativas a calidad y facturación, selecciona aquella RAN que ofrece el servicio requerido con el nivel de calidad deseado y con el menor coste. El valor del coste difundido por cada RAN no corresponde al coste económico derivado de la prestación del servicio, sino que se obtiene como el cociente entre la cantidad de recursos que se estima que va a consumir el usuario en caso de ser admitido (para cada servicio y cada nivel de calidad de servicio) y la cantidad de recursos disponibles, los cuales dependen de la carga de usuarios. El objetivo de esta definición de coste es el de que cada usuario se mantenga asignado a aquella RAN en la que provoca el menor consumo de recursos y, por tanto, el menor descenso en la capacidad disponible, en términos de número de

usuarios que pueden ser aceptados. En el artículo se proponen diversas variantes. En primer lugar distingue entre un *algoritmo básico*, consistente en aplicar el criterio de decisión anterior, y un *algoritmo con reselección* en el cual se contempla la posibilidad de que usuarios de UMTS que experimentan condiciones de transmisión radio poco favorables sean traspasados a GPRS. En segundo lugar, se establecen dos variantes adicionales en función de si el coste difundido por cada RAN se calcula a nivel de RAN (a partir de la información correspondiente a toda la RAN) o a nivel de celda (a partir de la información correspondiente a cada celda). Los resultados muestran que las mejores prestaciones corresponden al algoritmo con reselección basado en cálculos a nivel de celda. Una característica interesante de esta propuesta es que la selección de RAN se realiza por parte del terminal móvil sin necesidad de que las diferentes RANs deban intercambiar información entre sí. No obstante, sí se hace necesario acordar previamente un criterio común para determinar los niveles de calidad de servicio y los costes asignados a cada servicio. Este planteamiento resulta compatible con un escenario en el cual las diversas RANs son operadas por diferentes compañías y, por lo tanto, ninguna de ellas conoce las condiciones de operación de las demás.

Dentro del conjunto de propuestas que se caracterizan por el uso de funciones objetivo, puede identificarse un subconjunto en el cual el problema de la distribución de tráfico en redes heterogéneas se plantea como un problema de optimización, siendo la solución o distribución de tráfico óptima aquella que maximiza el valor de la función objetivo propuesta. Un ejemplo de este tipo de planteamiento es el trabajo presentado en [71, 72], donde se describe un criterio para decidir la distribución óptima, entre diferentes RANs, de un determinado volumen de demanda de tráfico. Dicho criterio aparece enmarcado dentro de una arquitectura CRRM específica y considerando un modelo de negocio particular. El modelo de negocio considerado contempla la existencia de diferentes operadores o proveedores de red independientes entre sí pero que mantienen cierta relación contractual entre ellos para poder compartir el tráfico de sus respectivas RANs. El artículo propone una arquitectura CRRM adecuada para dar soporte a dicho modelo de negocio, describiendo el proceso de comunicación entre sus elementos y proponiendo un criterio para decidir la distribución óptima del tráfico. Concretamente, el proceso se inicia cuando un determinado proveedor de red recibe una nueva solicitud de servicio o detecta una variación en las condiciones de funcionamiento de la red, momento en el cual puede decidir realizar una solicitud a los proveedores de red asociados. Los proveedores asociados procesan la solicitud y devuelven al proveedor de origen una respuesta en la cual especifican una determinada oferta. Los elementos de dicha oferta son dos. El primero determina el coste económico que impone el proveedor asociado por cada unidad de volumen de demanda, especificando un valor para cada tipo de servicio y para cada nivel de calidad con que puede ofrecerse. El segundo elemento de la oferta es el rango de volumen de demanda (expresado por ejemplo en número de usuarios, tasa binaria o Erlangs) para el cual la oferta realizada es válida (el trabajo no especifica ningún criterio o procedimiento para elaborar dicha oferta). La información obtenida de las ofertas de los proveedores asociados es procesada por el proveedor de origen junto con información proveniente de otras dos fuentes diferentes. La primera de estas fuentes es la nueva solicitud de servicio o variación detectada en las condiciones de funcionamiento de la red, a partir de la cual se especifica el servicio o conjunto de

servicios involucrados, las clases o tipos de usuarios de dicho servicio, los niveles de calidad de servicio permitidos para cada clase de usuario así como información relativa al perfil de usuario y las capacidades de su terminal, el volumen de demanda para cada clase de usuario, de servicio y de calidad de serivicio y, por último, información relativa a la zona horaria. La segunda de estas fuentes de información es el estado de la red, de donde se determina el volumen de demanda y la calidad de servicio ofrecida (por servicio y clase de usuario), la capacidad disponible y la utilidad derivada de prestar los servicios ofrecidos a cada nivel de calidad de servicio permitido. Los parámetros de estas tres fuentes de información (ofertas de los proveedores asociados, solicitud del servicio y estado de la red) aparecen reflejados en una función objetivo. El proveedor de red debe decidir una distribución del tráfico entre su propia RAN y las RANs de los proveedores asociados de manera que se maximice el valor de la función objetivo considerada. La optimización de esta función está sujeta a ciertas restricciones impuestas por el perfil de cada usuario, las prestaciones de su terminal móvil o el rango de validez de cada oferta, entre otros aspectos. Para encontrar la solución óptima a dicho problema, el artículo discute diversos algoritmos de optimización computacionalmente eficientes que permiten la obtención de una solución *quasi-óptima* en un tiempo razonable.

En [128, 129] el problema de la distribución de tráfico también se plantea como un problema de optimización, aunque con un enfoque distinto. En este caso la función objetivo representa una estimación de la capacidad global de la red heterogénea a partir de diversos parámetros. Por lo tanto, el objetivo del problema planteado es hallar una distribución de tráfico que maximice la capacidad global del sistema. Mediante la aplicación de técnicas de optimización matemática, se obtiene una solución *quasi-óptima* que lleva a la conclusión de que la distribución de servicios entre las RANs del sistema que maximiza la capacidad global depende de las curvas de capacidad de cada RAN. En unos casos la solución óptima consiste en mezclar, según cierta distribución, los diferentes tipos de servicio en cada una de las RANs disponibles, mientras que en otros casos la solución óptima consiste en aislar completamente los servicios, es decir, asignar cada tipo de servicio a una RAN diferente. Basándose en las conclusiones obtenidas, los autores proponen un sencillo algoritmo consistente en medir la distribución de servicios entre las diferentes RANs y realizar la asignación de los usuarios de forma que la distribución resultante se mantenga próxima a la distribución óptima obtenida al resolver matemáticamente el problema de optimización.

Para concluir esta sección dedicada a revisar las soluciones propuestas en la literatura por otros autores, citaremos un último enfoque, distinto a todos los anteriores, para resolver el problema de distribuir el tráfico de una red heterogénea entre sus diferentes RANs. Este enfoque se basa en el concepto de *radio multihoming* [130], según el cual un mismo terminal móvil mantiene varias conexiones o enlaces radio de manera simultánea con diversas RANs de una misma red heterogénea. La referencia [55] constituye un ejemplo de aplicación de este concepto a la distribución de tráfico en redes heterogéneas. Considerando una red heterogénea formada por una red celular (UMTS) y una red WLAN (HIPERLAN/2), propone dividir el tráfico a transmitir en dos flujos de información, uno con información básica y otro con información adicional. Cada uno de estos flujos de información se transmiten simultáneamente de forma paralela

a través de las dos RANs disponibles en el escenario considerado. Concretamente, el criterio propuesto consiste en transmitir el flujo de información básica por la red celular y el flujo de información adicional por la red WLAN. Este reparto se realiza atendiendo a las características y restricciones propias de cada RAN ya que, por lo general, las redes WLAN ofrecen elevadas tasas binarias de transmisión aunque con una conectividad limitada (pues suelen centrarse principalmente en *hot spots*), mientras que las redes celulares ofrecen una elevada conectividad pero generalmente con menores prestaciones binarias (por ejemplo, una portadora de 16 kb/s está disponible en el 99 % de los casos [55]). Dada la elevada disponibilidad de la red celular, este mecanismo permite garantizar al usuario una calidad de servicio mínima que vendría determinada por el flujo de información básica recibido a través de la red celular. Además, en aquellas localizaciones geográficas donde esté disponible una red WLAN y, por tanto, una mayor velocidad binaria de transmisión, el usuario recibiría también el flujo de información adicional que, combinado con el flujo de información básica, permitiría al usuario obtener una mayor calidad de servicio. Este esquema requiere en receptor una buena sincronización y multiplexación de los dos flujos de información asociados que se transmiten por RANs diferentes, lo cual posee un nivel de complejidad bastante considerable debido a que los retardos de cada red son, por lo general, bastante diferentes, imponiendo así problemas tanto de retardo como de *jitter*. Este esquema requiere además de una política de *scheduling* eficiente que permita compensar, en la medida de lo posible, estos inconvenientes. En [131] se discuten en detalle diversos algoritmos de *scheduling* así como varias técnicas de sincronización apropiadas para esta propuesta.

Un último ejemplo de aplicación del concepto de *radio multihoming* al problema que concierne al presente trabajo es el presentado en [132]. En este artículo se proponen y comparan dos esquemas de distribución de tráfico pensados para ser aplicados en un escenario en el cual los terminales móviles mantienen varias conexiones simultáneas según el principio de *multihoming* y además viajan en un medio de transporte a través de una red heterogénea. Durante el trayecto, los usuarios experimentan cambios en la red, atravesando regiones donde el conjunto de tecnologías de acceso radio que ofrecen cobertura varía frecuentemente a lo largo del recorrido. Los esquemas de distribución de tráfico propuestos pretenden determinar la manera óptima de distribuir el tráfico de datos de un usuario dado entre las diferentes conexiones radio simultáneas establecidas con la red a través de diferentes RANs. Un primer esquema denominado *in-transit* supone un desconocimiento total de los posibles cambios de red que pueden ocurrir durante el viaje y calcula la distribución de tráfico óptima sobre la marcha, cada vez que se detecta un cambio en la red. Un segundo esquema, denominado *pre-transit* supone un conocimiento *a priori* de los potenciales cambios de la red en un trayecto dado (suposición aplicable en el caso de transportes públicos con rutas repetitivas) y calcula todas las posibles distribuciones de tráfico antes de que dé comienzo el viaje. En ambos esquemas, el objetivo de los cálculos realizados es el de obtener una distribución de tráfico óptima que maximice el beneficio económico del operador de red, según cierto modelo de tarificación descrito en el artículo, al mismo tiempo que proporciona al usuario una calidad de servicio aceptable. Este trabajo constituye una extensión del trabajo realizado en [133], donde se realiza un estudio similar para el caso de usuarios con baja movilidad (*multihoming* estático).

2.2. Proyectos de investigación y estandarización

Varios de los trabajos presentados en la sección 2.1 fueron desarrollados bajo diversos proyectos de investigación IST (*Information Society Technologies*) de la Unión Europea. En esta sección se presentan y describen brevemente algunos de estos proyectos así como diversas actividades de estandarización relacionadas con este trabajo.

Uno de los principales objetivos del proyecto DRiVE (*Dynamic Radio for IP-Services in Vehicular Environments*) [134] y de su sucesor OverDRiVE (*Spectrum Efficient Uni- and Multicast Services over Dynamic Multi-Radio Networks in Vehicular Environments*) [135] es el de proveer servicios multimedia inalámbricos de alta calidad de manera espectralmente eficiente en redes heterogéneas, centrando la atención sobre la prestación de dichos servicios a vehículos. Uno de los principales retos para alcanzar este objetivo es el del lograr un buen interfuncionamiento de diferentes sistemas radio como GERAN, UTRAN, DAB y DVB-T dentro de una banda de frecuencias común. Los aspectos RRM cubiertos dentro de estos dos proyectos son la mejora de la eficiencia espectral mediante técnicas de asignación dinámica de espectro (DSA) y algoritmos avanzados para la selección de RAN, caso que corresponde al presente trabajo. En [136, 137] se describen los algoritmos DSA diseñados e investigados para una red formada por UMTS y DVB-T. Los aspectos relativos a la distribución de tráfico en estos proyectos puede encontrarse en [138, 139]. En [138] se describe la función de control de tráfico en el proyecto DRiVE desde los puntos de vista funcional y estructural. En [139] se describe el trabajo realizado relativo a la distribución del tráfico de la red sobre las diferentes RANs de una red heterogénea. En las propuestas estudiadas, es el terminal móvil el que toma la decisión de cuál va a ser la RAN a través de la cual va a recibir el servicio. Estas decisiones se toman a partir de la información generada y difundida por cada RAN del sistema. Los algoritmos analizados se centran sobre dos escenarios diferentes. El primer escenario considerado comprende dos sistemas celulares (UMTS y GSM/GPRS), situación bastante realista que cabe esperar en el desarrollo hacia la 4G. El segundo escenario incluye una red celular (UMTS) y una red de difusión (DVB-T), situación interesante en la cual se estudia la cooperación de sistemas que normalmente son operados de forma completamente independiente.

El objetivo del proyecto EVEREST (*Evolutionary Strategies for Radio Resource Management in Cellular Heterogeneous Networks*) [140] es el desarrollo y evaluación de un conjunto específico de estrategias y algoritmos, tanto para la red de acceso como para la red troncal, que permitan la prestación de diferentes servicios en redes heterogéneas con garantías de calidad extremo a extremo y de una utilización óptima de los escasos recursos radio disponibles. Las principales tareas llevadas a cabo durante el proyecto fueron, entre otras, las de elaborar algoritmos RRM avanzados para GERAN y UTRAN, así como algoritmos CRRM para redes heterogéneas formadas por GERAN y UTRAN. En este proyecto también se consideraron otras tecnologías complementarias a GERAN y UTRAN como, por ejemplo, WLAN para *hot spots* en entornos interiores. La parte del proyecto relacionada con este trabajo puede encontrarse en [141], donde se proporciona la definición de una primera versión de los algoritmos RRM/CRRM considerados, y en [142], donde se evalúan los algoritmos seleccionados en [141].

Otros proyectos europeos que también guardan cierta relación con este trabajo, aunque de forma menos directa, son los proyectos *Ambient Networks* [143], MONASIDRE (*Management Of Networks And Services In a Diversified Radio Environment*) [144] y WINNER (*Wireless World Initiative New Radio*) [145].

El objetivo del proyecto *Ambient Networks*, en lo que a RMM/CRRM se refiere, es el diseño de nuevos algoritmos para la gestión eficiente de los recursos radio a través de diferentes RATs y múltiples operadores para proveer un acceso eficiente y de bajo coste, independientemente de las fronteras administrativas.

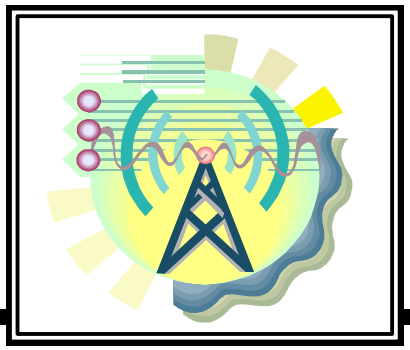
El objetivo del proyecto MONASIDRE es la implementación de una sofisticada plataforma para la gestión de redes y servicios en una red heterogénea integrada por UMTS, sistemas de difusión digital (DBSs) y sistemas móviles de banda ancha (*Mobile Broadband Systems*, MBSs).

El proyecto WINNER considera algunos aspectos RRM en redes con múltiples alternativas de acceso radio, aunque su principal objetivo es el diseño de un único sistema de acceso radio, escalable y adaptable, que permita acceso ubicuo en escenarios tanto de corto como de largo alcance. Los aspectos CRRM considerados en este proyecto son la definición de un mecanismo de cooperación a nivel radio entre diferentes RANs y la elaboración de métodos para un uso flexible y eficiente del espectro.

En cuanto a actividades de estandarización relacionadas con el presente trabajo, el proyecto 3GPP (*3rd Generation Partnership Project*) ha estudiado diferentes posibilidades de interfuncionamiento entre redes GSM y UMTS. Dentro de este proyecto de estandarización se han llevado a cabo diversos estudios relativos a la distribución de tráfico entre GERAN y UTRAN, los cuales se describen en [146, 147]. En estos informes puede encontrarse un estudio detallado de la propuesta realizada en [120, 121]. Para otras RATs como WLAN o Bluetooth, 3GPP ha dado ya los primeros pasos hacia la coordinación de estas redes entre sí y con redes celulares ya existentes [148–151].

2.3. Sumario

En este capítulo se ha realizado una revisión del trabajo previo de otros autores en el campo de la distribución de tráfico en redes heterogéneas. En la primera parte del capítulo se han descrito varias propuestas realizadas por otros autores en la literatura científico-técnica. Inicialmente se han descrito varias soluciones sencillas, introduciendo así la problemática asociada al presente trabajo, para a continuación presentar de forma progresiva otras propuestas más sofisticadas y complejas. Las soluciones descritas, a menudo elaboradas de forma independiente e inconexa, han sido presentadas procurando mantener una agrupación lógica según diferentes enfoques adoptados a la hora de resolver el problema. En la segunda parte del capítulo se han citado varios proyectos de investigación europeos relacionados con el presente trabajo, así como diversos estudios de interés llevados a cabo en el ámbito de la estandarización.



CAPÍTULO 3

Algoritmos Propuestos

En este capítulo se presentan y describen nuevos algoritmos CRRM diseñados para distribuir de forma eficiente el tráfico de una red radio heterogénea entre las diversas alternativas de acceso radio (RATs) disponibles. El problema que se plantea consiste en decidir, para cada usuario del sistema, cuál de las RATs disponibles debe ser seleccionada para proveer el servicio solicitado, de tal manera que los requisitos de calidad de servicio se vean satisfechos al mismo tiempo que se optimiza la utilización global de los recursos radio del sistema. Cada uno de los algoritmos propuestos se basa en un criterio diferente [152]. No obstante, todos ellos tienen en común el empleo de una determinada función de utilidad que cuantifica numéricamente todos los factores tenidos en cuenta por el criterio de decisión aplicado. En el primero de los algoritmos propuestos la función de utilidad proporciona, a partir de la calidad radio y de la disponibilidad de recursos, una estimación del *throughput* que cabría esperar de cada una de las RATs del sistema. Por ello, este algoritmo será denominado en lo sucesivo algoritmo UBET (*Utility Based on Expected Throughput*). En el segundo algoritmo propuesto la tasa de transmisión también es tenida en cuenta, pero la utilidad que un usuario dado asigna a una determinada tasa de transmisión depende de la requerida por el servicio solicitado para ser provisto satisfactoriamente. Así pues, este algoritmo será denominado en lo sucesivo algoritmo UBReQoS (*Utility Based on Required Quality of Service*). Los algoritmos propuestos en este capítulo han sido diseñados considerando una red heterogénea genérica. No obstante, la definición de los mismos se ilustrará considerando una red integrada por GPRS, EDGE y HSDPA, ya que estas tres RATs son las implementadas en el simulador utilizado en este trabajo. Se recomienda una lectura previa del apéndice A, donde se puede encontrar una descripción de estas tres tecnologías.

3.1. Algoritmo UBET

Una red radio heterogénea se caracteriza por la presencia de diversas RATs, cada una de las cuales ofrece por lo general distintas prestaciones, tal y como se mostró en la tabla 1.1 del capítulo 1. Como primera aproximación, podría proponerse que siempre se seleccione aquella RAT que ofrezca las mejores prestaciones, por ejemplo, en términos de velocidad de transmisión de datos. De esta forma cabría esperar, en principio, la obtención de elevados niveles de calidad de servicio (*Quality of Service*, QoS) para los usuarios del sistema. Sin embargo, es necesario tener en cuenta dos aspectos. En primer lugar, según se comentó en el capítulo 1, esta solución podría dar lugar a una situación indeseable en la cual una de las RATs del sistema permanece fuertemente cargada, incluso saturada, mientras los recursos de otras RATs disponibles en el sistema están ociosos, es decir, sin utilizar. Esta observación sugiere la necesidad de distribuir los usuarios de una red heterogénea entre las RATs disponibles con la finalidad de conseguir una utilización más eficiente de los recursos del sistema. En segundo lugar, la estrategia anterior permitiría la obtención de elevados niveles de QoS mientras la carga del sistema se mantuviera en niveles bajos o moderados. Si la carga del sistema aumenta y todos los usuarios del sistema son siempre dirigidos hacia la RAT con mejores prestaciones, la disponibilidad de recursos en dicha RAT decrece al mismo tiempo que los niveles de interferencia experimentados aumentan. En esta situación, el rendimiento real de dicha RAT se vería reducido y podría darse el caso de que otra RAT alternativa menos cargada ofreciera un mejor rendimiento, aun cuando las velocidades máximas de transmisión de dicha RAT fueran de hecho menores.

En este contexto se propone el algoritmo UBET, cuyo objetivo es el de ofrecer elevados niveles de QoS a los usuarios del sistema dirigiéndolos preferiblemente hacia la RAT del sistema con mayores prestaciones, hasta que la carga en dicha RAT alcance un nivel lo suficientemente elevado como para que otras RATs alternativas del sistema ofrezcan un rendimiento previsiblemente mejor, en cuyo caso los nuevos usuarios son asignados a dichas RATs alternativas. Para tal fin se propone el uso de un conjunto de funciones de utilidad, una para cada RAT del sistema, que asignan un valor de utilidad a cada una de las RATs disponibles. Dicho valor de utilidad representa una estimación del *throughput* que cabe esperar en cada RAT si ésta es seleccionada. Esta estimación se realiza a partir de la calidad de transmisión radio experimentada y de la disponibilidad de recursos en la RAT considerada. Cada vez que sea necesario tomar una decisión, la RAT seleccionada será aquella que ofrezca el mayor valor de utilidad.

La función de utilidad U de cada RAT se define como el producto de otras dos funciones de utilidad U_1 y U_2. La función de utilidad U_1 ofrece una estimación del *throughput* basada únicamente en la calidad de transmisión radio experimentada. Para obtener el valor de la función U_1 de una RAT determinada, se emplea un conjunto de curvas que representan, para cada uno de los modos de transmisión (modulación y codificación de canal) disponibles en dicha RAT, la tasa de errores de transmisión, expresada por ejemplo en términos de la proporción de bloques erróneos (*BLock Error Rate*, BLER), en función de la calidad experimentada en el enlace radio, expresada por ejemplo en términos de la relación portadora a interferencia (*Carrier-to-Interference*

Ratio, CIR) o señal a interferencia (*Signal-to-Interference Ratio*, SIR)[1]. En las figuras A.2, A.7 y A.10 del apéndice A se muestra un ejemplo de este tipo de curvas para los modos de transmisión de GPRS, EDGE y HSDPA. A partir de estas curvas, el *throughput* puede expresarse entonces como

$$Throughput = R \cdot (1 - BLER) \tag{3.1}$$

donde R corresponde a la tasa binaria del modo de transmisión considerado (ver tablas A.1, A.2 y A.3). Las figuras A.3, A.8 y A.11 muestran un ejemplo del valor del *throughput* en función de la calidad del enlace radio para los modos de transmisión de GPRS, EDGE y HSDPA. Estas curvas han sido obtenidas aplicando la expresión 3.1 a las curvas representadas en las figuras A.2, A.7 y A.10. Suponiendo que en todo momento se selecciona el modo de transmisión que ofrece el máximo valor de *throughput* para la calidad experimentada en el enlace radio, la envolvente de estas curvas puede entonces utilizarse como una estimación del *throughput* en función de la calidad radio si la RAT considerada es seleccionada. Así pues, esta envolvente se toma como función de utilidad U_1. Las figuras 3.1, 3.2 y 3.3 muestran cómo se obtiene la curva correspondiente a la función de utilidad U_1 para GPRS, EDGE y HSDPA.

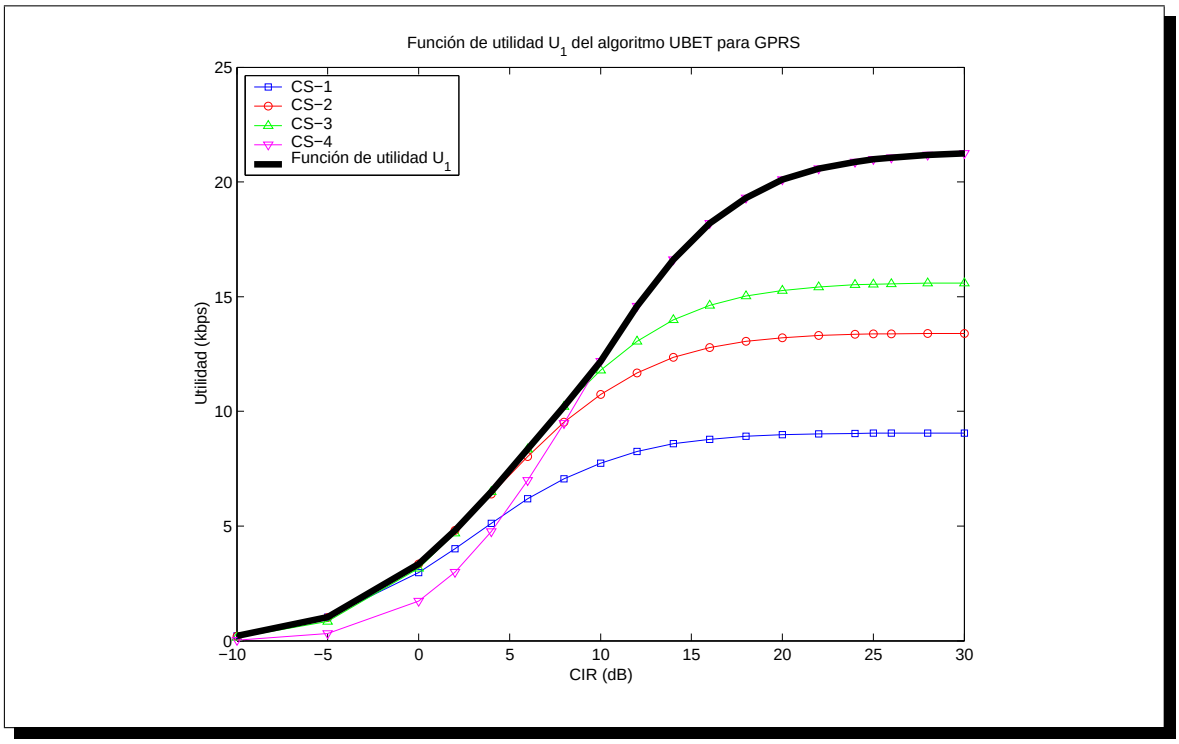


Figura 3.1: Función de utilidad U_1 del algoritmo UBET para GPRS.

[1]En este trabajo se emplea el término CIR para GPRS y EDGE, ya que en dichos sistemas, basados en un acceso múltiple de tipo TDMA/FDMA, existe cierta identificación entre el nivel de señal útil y el nivel de señal portadora recibido. Para HSDPA, basado en un acceso múltiple de tipo CDMA, se ha preferido emplear el término SIR con el fin de enfatizar el hecho de que el nivel de señal útil corresponde a una parte de la potencia de la frecuencia portadora.

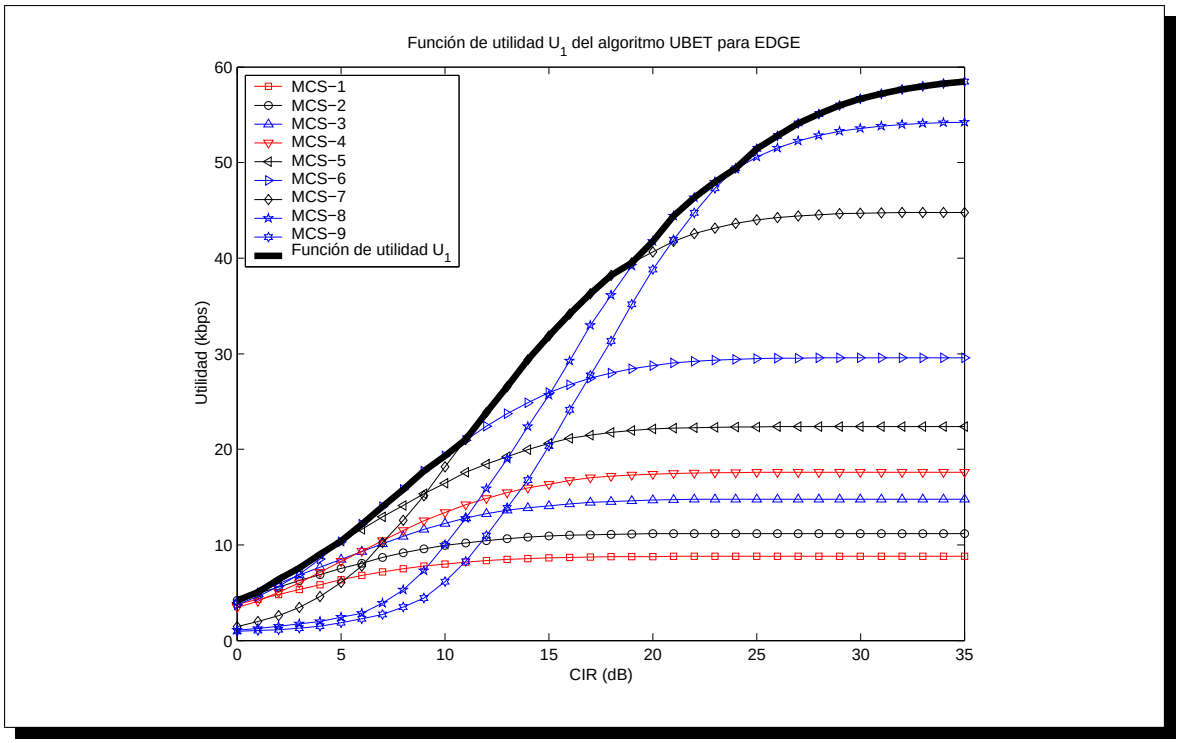


Figura 3.2: Función de utilidad U_1 del algoritmo UBET para EDGE.

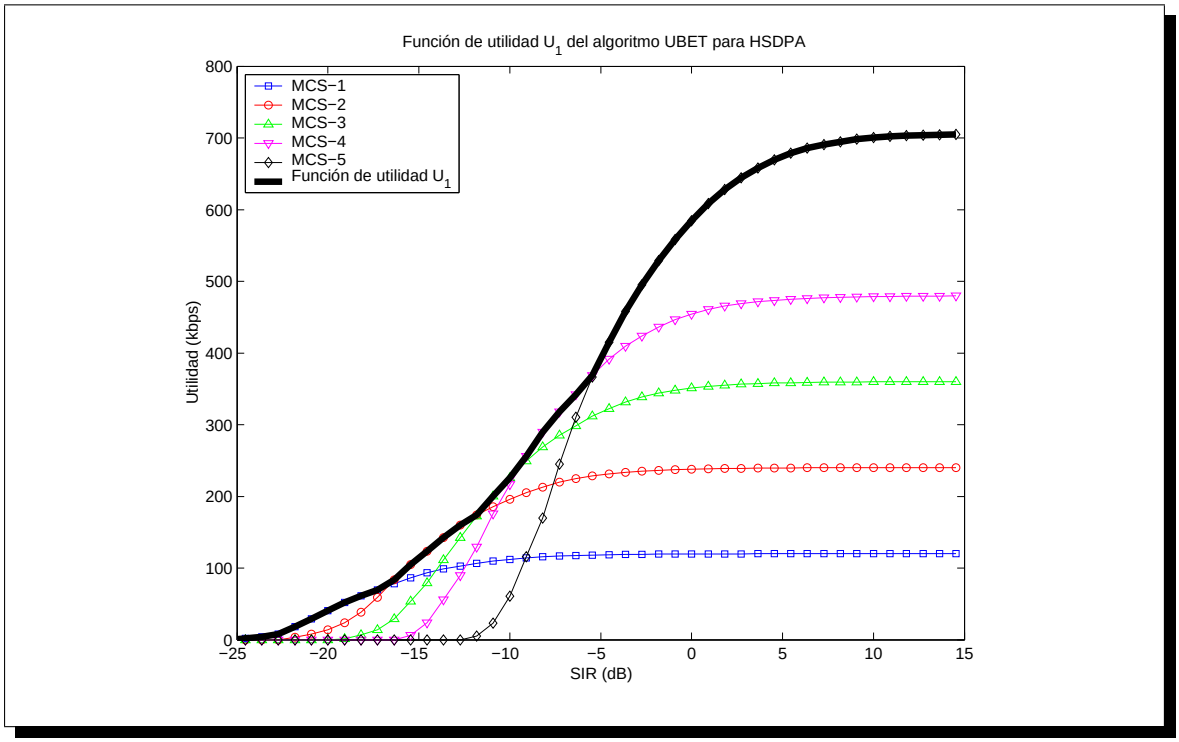


Figura 3.3: Función de utilidad U_1 del algoritmo UBET para HSDPA.

Puesto que el valor de la función U_1 de una RAT dada está directamente relacionado con las velocidades de transmisión de dicha RAT, cabe esperar que la RAT con mayores velocidades sea seleccionada habitualmente en primer lugar (en el caso particular considerado, sería HSDPA). No obstante, si el número de usuarios en dicha RAT aumenta, los niveles de interferencia aumentarán también. Este hecho dará como resultado un deterioro en la calidad de transmisión radio experimentada (reducción del valor de CIR/SIR), lo cual a su vez ocasionará una reducción del *throughput* estimado que cabe esperar para los usuarios de dicha RAT. Si la carga aumenta lo suficiente, una RAT alternativa con una carga menor y, por tanto, con menores niveles de interferencia (valores mayores de CIR/SIR) podría ofrecer un mejor rendimiento en términos de *throughput*. En este caso el valor de utilidad U_1 sería mayor para la RAT alternativa menos cargada. Consecuentemente, los nuevos usuarios serían asignados a dicha RAT, donde obtendrían un mayor *throughput*. De esta forma se mejoraría el nivel de QoS global al mismo tiempo que se prevendría la saturación en la RAT con mejores prestaciones. Si las prestaciones ofrecidas por las RATs del sistema difieren notablemente, cabe la posibilidad de que la RAT con mayores velocidades siga ofreciendo el mayor valor de *throughput* esperado incluso ante elevadas cargas. En este caso puede ser conveniente que las RATs con menores velocidades asignen simultáneamente varios canales a cada usuario para así poder ofrecer velocidades de transmisión comparables.

La función de utilidad U_1 ofrece una estimación del *throughput* que puede esperarse de una determinada RAT. Esta estimación tiene en cuenta el efecto de los usuarios en celdas interferentes a través del valor de la relación CIR/SIR. Sin embargo, es importante observar que cuando se calcula el valor de U_1 para decidir a qué RAT asignar un usuario determinado, no se tiene en cuenta el efecto del resto de usuarios de su misma celda que se hallan conectados a la misma RAT y que, por lo tanto, compiten por los canales disponibles. Si el número de usuarios de la misma celda es lo suficientemente grande en comparación con el número de canales disponibles para la RAT considerada, no todos los usuarios recibirán un canal cada vez que lo soliciten, en cuyo caso habrán de esperar en una cola, lo cual repercutirá sobre el nivel de QoS percibido por el usuario. Con la finalidad de incluir este efecto en el valor de utilidad de cada RAT, se introduce una segunda función de utilidad U_2, definida como

$$U_2 = \begin{cases} 1, & N_u \leq N_c \\ N_c/N_u, & N_u > N_c \end{cases} \tag{3.2}$$

donde N_u representa el número de usuarios en la misma celda que se encuentran conectados a la misma RAT y N_c representa el número total de canales para dicha RAT en dicha celda. La figura 3.4 muestra el valor de U_2 en función de N_u. Según se aprecia, si el número de usuarios es menor o igual que el número de canales disponibles, puede garantizarse que un usuario que solicite un canal libre siempre lo recibirá. Por lo tanto, para $N_u \leq N_c$, $U_2 = 1$ y $U = U_1$, lo cual significa que el *throughput* estimado viene dado por el valor de U_1. Sin embargo, si el número de usuarios es mayor que el número de canales, algunos usuarios podrían no recibir un canal libre siempre que lo solicitaran. Por ejemplo, si hay el doble de usuarios que de canales, es decir, $N_u = 2 \cdot N_c$, puede considerarse como aproximación sencilla que, en promedio, cada usuario tendría acceso

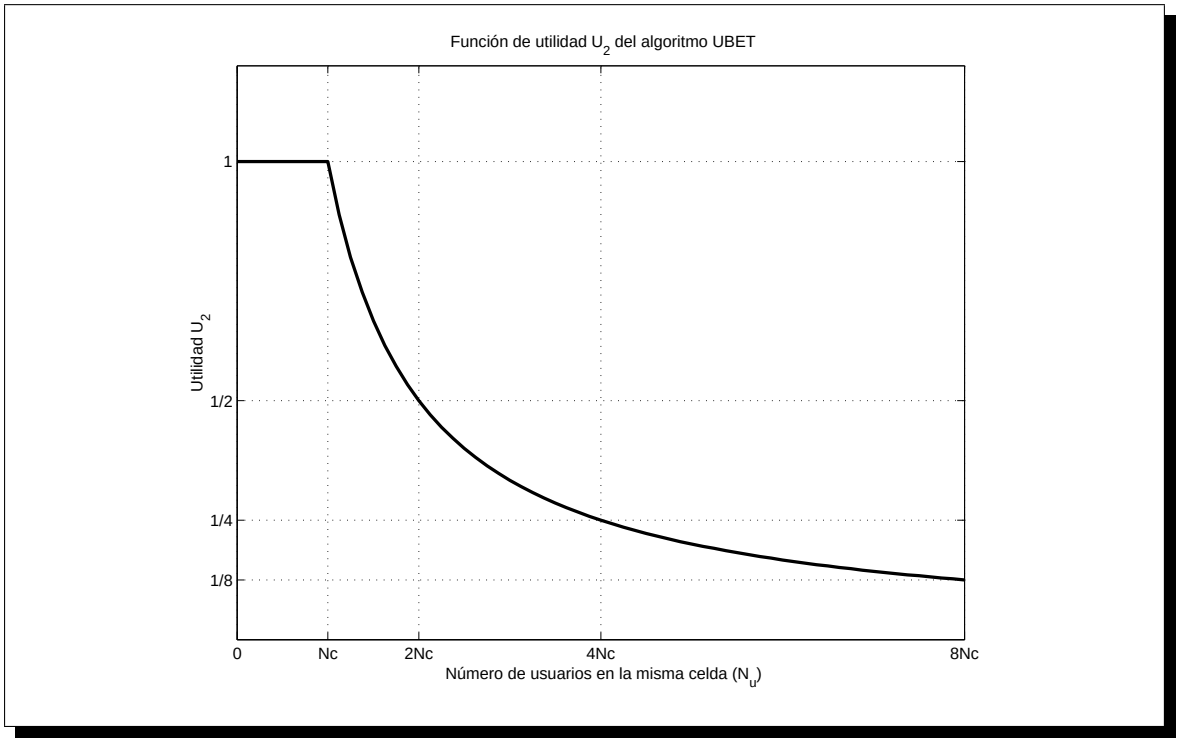


Figura 3.4: Función de utilidad U_2 del algoritmo UBET.

a un canal de transmisión la mitad de las veces que lo solicitara (en realidad, esto podría depender de la política de *scheduling* aplicada). Por lo tanto, para $N_u = 2 \cdot N_c$, $U_2 = 0{,}5$ y $U = 0{,}5 \cdot U_1$, lo que significa que el *throughput* esperado en este caso se supone aproximadamente la mitad de la estimación dada por U_1.

El valor de la función de utilidad U_2 para una RAT dada puede obtenerse fácilmente ya que depende únicamente del número de usuarios que se hallan conectados a dicha RAT en la celda considerada. Sin embargo, el cálculo del valor de la función de utilidad U_1 no resulta tan sencillo, pues para ello primero debe estimarse el nivel de interferencia experimentado en cada una de las RATs del sistema a fin de obtener un valor de la relación CIR/SIR. Esta operación, ciertamente compleja en un sistema real, debe realizarse además cada vez que sea necesario decidir a qué RAT debe asignarse un usuario dado. Con el fin de simplificar este procedimiento, se propone una aproximación que puede emplearse para determinar de manera sencilla un valor de CIR/SIR representativo de la calidad de transmisión radio en cada una de las RATs del sistema. El método sugerido ha sido elaborado considerando el enlace descendente y se basa en los niveles de señal predichos por un determinado modelo de propagación. En este trabajo se ha considerado el modelo Okumura-Hata COST 231, cuya descripción se encuentra en la sección 4.2.2, aunque cabe la posibilidad de emplear otro modelo distinto si éste resultara más apropiado al entorno de propagación en el cual se pretenda aplicar el algoritmo. El método propuesto sugiere calcular por separado el nivel de señal/portadora y el nivel de interferencia, suponiendo que transmisor y receptor se encuentran en unas localizaciones determinadas dentro de sus respectivas celdas. En primer lugar, para obtener el nivel de señal/portadora se supone que el terminal móvil

se encuentra justo en el borde de la celda y que la estación base se halla en el centro de dicha celda. Esta distribución, ilustrada en la figura 3.5, corresponde al caso más desfavorable ya que en esta situación el nivel de señal recibido es mínimo. Con esta manera de proceder se pretende en cierta forma que la velocidad de transmisión de datos experimentada en la RAT que sea seleccionada no sea inferior a la estimación predicha por la función de utilidad del algoritmo a partir de la calidad del enlace radio. Así pues, suponiendo esta separación entre estación base y terminal móvil, se calcula el nivel de señal/portadora recibido por el móvil haciendo uso del modelo de propagación considerado. Para calcular el nivel de interferencia se tiene en cuenta el nivel de señal recibido de las celdas pertenecientes a la primera corona de celdas interferentes (en la sección 4.2.4.2 se explica cómo se determinan estas celdas). La interferencia de la segunda corona y posteriores es considerablemente menor y puede omitirse. Para el cálculo del nivel de interferencia se realiza una distinción entre diferentes sistemas.

En sistemas FDMA/TDMA, como es el caso por ejemplo de GPRS y EDGE, la interferencia procede de celdas cocanales separadas de la celda interferida por una distancia igual a la distancia de reutilización cocanal (considerando sólo la primera corona de celdas interferentes). En este caso suponemos que tanto las estaciones base interferentes como el terminal móvil interferido se encuentran en el centro de sus respectivas celdas, tal y como se muestra en la figura 3.6. La finalidad de considerar que el terminal móvil interferido se halla en el centro de su celda es la de eliminar los efectos que tiene la movilidad del usuario sobre el valor instantáneo de la calidad de transmisión experimentada en el enlace radio. Con todo ello, el valor de la relación CIR para GPRS y EDGE se obtendría como

$$CIR_{GPRS/EDGE} = \frac{\frac{P_i}{L_P(r)}}{N_j \cdot \frac{P_j}{L_P(R)} + N_0 \cdot W} \tag{3.3}$$

donde P_i es la potencia de transmisión de la señal deseada en la celda de referencia (celda i), $L_P(r)$ son las pérdidas predichas por el modelo de propagación considerado entre el centro y el borde de la celda de referencia, es decir, sobre una distancia igual al radio r de la celda de referencia, N_j es el número de transmisores activos interferentes en celdas cocanales, P_j es la potencia de transmisión de las celdas interferentes (por simplicidad supuesta igual en todas las celdas interferentes), $L_P(R)$ son las pérdidas predichas por el modelo de propagación considerado sobre una distancia igual a la distancia de reutilización cocanal R, y $N_0 \cdot W$ representa el ruido térmico en el receptor, siendo N_0 la densidad espectral de ruido y W el ancho de banda del canal.

En el caso de sistemas CDMA, como es el caso de HSDPA, la interferencia procede de las celdas inmediatamente vecinas a la celda interferida (considerando la primera corona de celdas interferentes) pero también desde dentro de la propia celda. En efecto, aunque los códigos de canalización (ver apéndice A) de una misma celda son perfectamente ortogonales entre sí, el efecto multicamino reduce la ortogonalidad entre dichos

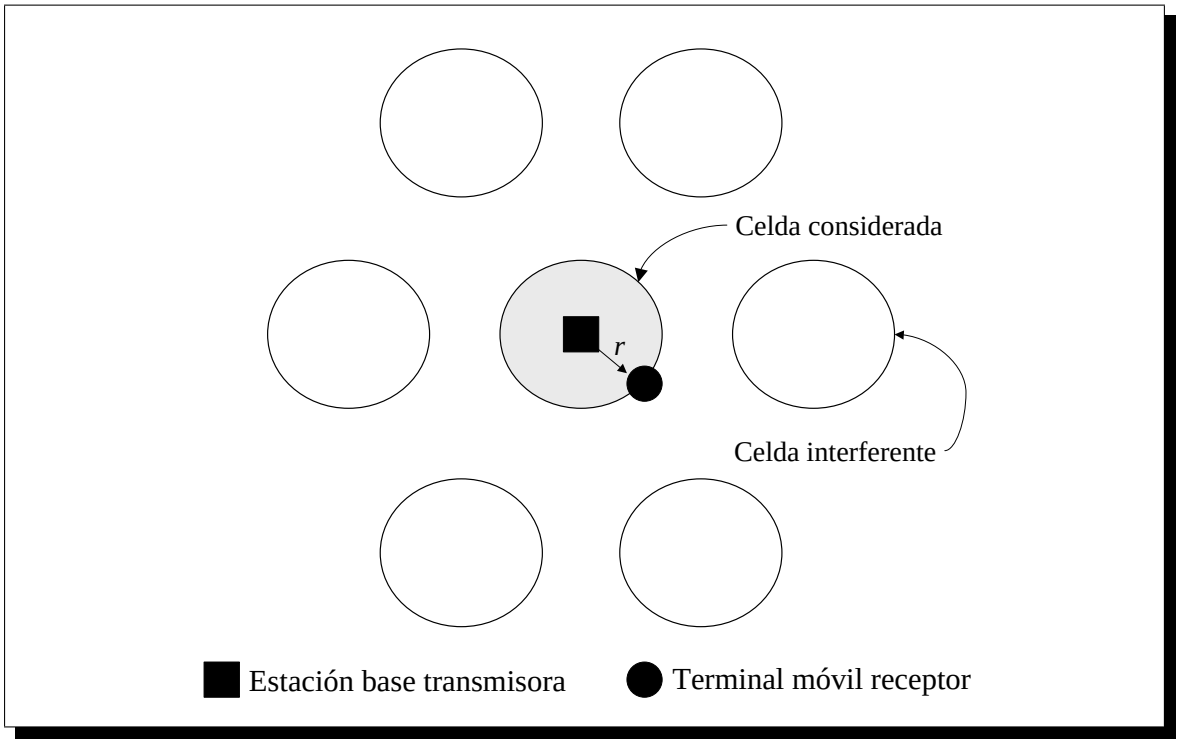


Figura 3.5: Aproximación para estimar el nivel de señal/portadora (en GPRS, EDGE y HSDPA) y de interferencia intracelda (sólo en HSDPA).

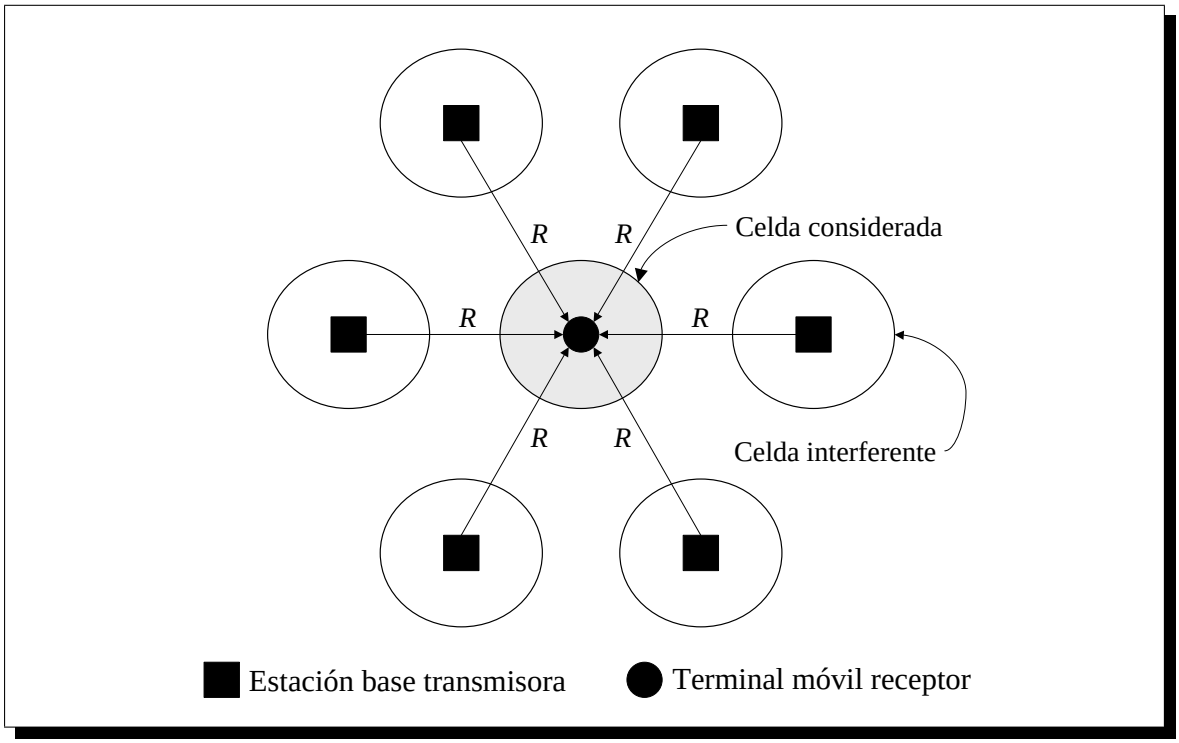


Figura 3.6: Aproximación para estimar el nivel de interferencia intercelda.

códigos, dando lugar a cierta componente de interferencia intracelda. Para obtener el nivel de interferencia intercelda se supone la misma distribución de estaciones base interferentes y usuario interferido que para sistemas FDMA/TDMA (ver figura 3.6). El nivel de interferencia intracelda se obtiene suponiendo que el usuario interferido se encuentra en el borde de la celda mientras que la interferencia del resto de usuarios de su misma celda procede de la estación base colocada en el centro de la celda, tal y como se muestra en la figura 3.5. En este caso, la potencia de transmisión de los usuarios que producen interferencia intracelda es atenuada por un factor de ortogonalidad α [153]. Por lo tanto, para HSDPA el valor de la relación SIR se calcula según indica la expresión 3.4, donde N_i representa el número de canales activos que están siendo utilizados por otros usuarios dentro de la misma celda.

$$SIR_{HSDPA} = \frac{\dfrac{P_i}{L_P(r)}}{N_i \cdot \dfrac{P_i \cdot (1-\alpha)}{L_P(r)} + N_j \cdot \dfrac{P_j}{L_P(R)} + N_0 \cdot W} \tag{3.4}$$

El método que se acaba de describir ofrece la posibilidad de establecer, de manera sencilla, una relación directa entre el número de usuarios interferentes y el valor de la relación CIR/SIR para cada una de las RATs del sistema. Esta información puede almacenarse en sencillas tablas, las cuales pueden ser consultadas en lugar de tener que estimar los niveles de interferencia durante la operación del sistema cada vez que sea necesario calcular el valor de la función de utilidad U_1 para tomar una decisión de selección de RAT. El método propuesto, así pues, supone una considerable reducción del coste computacional de implementación del algoritmo propuesto.

Resumiendo, el modo de operación del algoritmo UBET es el siguiente. Cada vez que es necesario decidir a qué RAT debe asignarse un usuario, se calcula el valor de la función de utilidad $U = U_1 \cdot U_2$ para cada una de las RATs disponibles en la posición geográfica del usuario. Para obtener el valor de la función de utilidad U_1 de una RAT determinada, puede estimarse el nivel de interferencia experimentado en dicha RAT y obtener así un valor para la relación CIR/SIR, o bien puede obtenerse el número de usuarios interferentes intercelda N_j e intracelda N_i (si procede) y, consultando una tabla almacenada, determinar el valor de relación CIR/SIR correspondiente. Los valores de dicha tabla habrán sido previamente calculados siguiendo el método anteriormente descrito. En cualquier caso, el valor de CIR/SIR obtenido por el procedimiento elegido se transforma a un valor de utilidad haciendo uso de la función de utilidad U_1, la cual ofrece una estimación del *throughput* basada en la calidad del enlace radio. Para calcular el valor de la función de utilidad U_2 de una RAT dada, simplemente es necesario contar el número de usuarios que se hayan conectados a dicha RAT en la misma celda en la cual se encuentra el usuario para el cual se va a decidir la RAT a asignar. Repitiendo este procedimiento para todas las RATs disponibles en la posición geográfica del usuario, se obtiene un valor de utilidad U para cada una de ellas. El usuario en cuestión será asignado a aquella RAT que ofrezca el valor de U más elevado.

3.2. Algoritmo UBReQoS

El criterio del algoritmo UBET consiste en asignar los usuarios del sistema a aquella RAT que previsiblemente ofrecerá en cada momento las mejores prestaciones en términos de *throughput*. Esta asignación se realiza ignorando la posibilidad de que los requisitos de QoS de algunos usuarios con servicios poco exigentes podrían verse satisfechos aun cuando dichos usuarios fueran asignados a otras RATs alternativas con prestaciones más modestas. Esta cuestión adquiere relevancia en escenarios en los cuales las RATs presentes en el sistema poseen prestaciones considerablemente diferentes, especialmente cuando los servicios ofrecidos en el sistema se caracterizan por requisitos de QoS bastante distintos y no todos ellos pueden satisfacerse en cualquiera de las RATs disponibles. En efecto, esto podría constituir un problema en aquellas ocasiones en las cuales los usuarios de servicios con bajos requisitos de QoS, para los cuales podría resultar suficiente una red de bajas prestaciones, acaparan la mayor parte de los recursos de las RATs con mejores prestaciones impidiendo de esta forma que otros usuarios de servicios con elevados requisitos de QoS y que, por tanto, necesitan imperativamente las RATs con mejores prestaciones, puedan recibir un servicio satisfactorio. Esta situación puede evitarse si los usuarios del sistema son asignados a una RAT cuyas prestaciones sean acordes al nivel de QoS requerido por el usuario. Como primera aproximación, podría asociarse cada servicio ofrecido en el sistema a una RAT cuyas capacidades permitieran satisfacer los niveles de QoS de dicho servicio. Así, cada vez que un usuario solicitara un determinado servicio, éste le sería siempre ofrecido a través de la RAT correspondiente. Este esquema, si bien resulta atractivo por su sencillez, posee el inconveniente de que un aumento en la demanda de un determinado servicio podría dar lugar a la saturación de la RAT asociada a dicho servicio, aun cuando otras RATs dispusieran de recursos libres. Por lo tanto, se hace necesario diseñar otros principios más elaborados que permitan satisfacer los distintos niveles de QoS requeridos haciendo un uso más eficiente de la totalidad de recursos radio disponibles en el sistema.

Con esta finalidad se propone el algoritmo UBReQoS, cuyo objetivo es el de distribuir inteligentemente los usuarios de una red radio heterogénea entre las RATs disponibles en función de las prestaciones de cada una de ellas y del nivel de QoS requerido por el servicio solicitado. Para tal fin, la función de utilidad definida para este algoritmo asigna a cada una de las RATs disponibles en el sistema un valor de utilidad que depende, no sólo de las capacidades de transmisión de dicha RAT como es el caso del algoritmo UBET, sino también de los requisitos de QoS específicos del servicio solicitado por cada usuario del sistema.

Para la aplicación del algoritmo UBReQoS, los requisitos de QoS de un determinado servicio se representan mediante la tasa binaria nominal R_{nom} requerida por dicho servicio para que pueda proveerse adecuadamente, de tal forma que la provisión de un servicio caracterizado por un determinado valor de R_{nom} se considera satisfactoria si la tasa binaria de transmisión experimentada es igual o mayor que R_{nom} y se considera no satisfactoria en caso contrario. Más adelante se discutirá en mayor detalle cómo se define dicho parámetro para diferentes tipos de servicio y cómo se establece su valor.

En la definición de la función de utilidad del algoritmo UBReQoS se introduce en primer lugar una función de utilidad u_1. La forma específica de esta función queda determinada por el valor del parámetro R_{nom} anteriormente citado (dependerá por tanto de los requisitos de QoS del servicio considerado) y por un nuevo parámetro que en lo sucesivo denominaremos *coeficiente de utilidad nominal* (CUN). El parámetro CUN se define como la utilidad que percibe el usuario cuando la tasa binaria de transmisión es igual a la tasa binaria R_{nom} requerida por el servicio que solicita dicho usuario. La figura 3.7 ilustra la definición de este parámetro. Considerando un servicio que requiere por ejemplo una tasa binaria $R_{nom} = 12$ kbps, un valor de CUN igual a 1.00 quiere decir que cuando la tasa binaria es de 12 kbps, el valor de utilidad u_1 percibido por el usuario del servicio es, precisamente, $u_1 = 1.00$. Este parámetro, por lo tanto, establece una relación entre la tasa binaria estimada que puede ofrecer el sistema y la tasa binaria R_{nom} requerida por el servicio en cuestión. Según se puede observar en la figura 3.7, la función de utilidad u_1 se define como una función de utilidad normalizada que toma valores dentro del intervalo $[0, 1]$. Para una tasa binaria estimada igual a cero, el valor de utilidad asociado es siempre igual a cero. A medida que la tasa binaria estimada aumenta, el valor de utilidad u_1 percibido por el usuario también aumenta hasta llegar al valor máximo, a partir del cual un aumento en la tasa binaria deja de producir un aumento en la utilidad percibida por el usuario. La pendiente de la recta en la región creciente viene determinada por el valor de R_{nom} que caracteriza el servicio considerado y por el valor del parámetro CUN empleado en el diseño de la función u_1.

El interés de definir la función de utilidad u_1 de la forma anteriormente indicada se ilustra en la figura 3.8. En dicha figura se muestra la forma de la función u_1 para distintos servicios caracterizados por diferentes valores de R_{nom}, es decir, por diferentes requisitos de QoS, considerando en todos ellos CUN = 1.0. Considérese por ejemplo un usuario de un servicio que requiere una tasa binaria R_{nom} relativamente baja o moderada. Para tasas binarias mayores que R_{nom} el usuario percibe la misma utilidad que para R_{nom}, lo cual significa que una RAT capaz de proveer una tasa binaria de transmisión al menos igual a R_{nom} resultará suficiente para satisfacer al usuario, por lo que no hay necesidad de asignar dicho usuario a otra RAT con mayores prestaciones binarias. Por lo tanto, dicho usuario será asignado a una RAT capaz de satisfacer sus requisitos de tasa binaria, aunque otras RATs del sistema ofrezcan mayores prestaciones. Por otra parte, el usuario de un servicio caracterizado por un valor relativamente alto de R_{nom} será asignado a las RATs con mejores prestaciones binarias, ya que otras RATs que ofrezcan tasas binarias inferiores a R_{nom} no serán capaces de satisfacer a dicho usuario. La función de utilidad u_1, así pues, ha sido diseñada con la intención de distribuir el tráfico de la red heterogénea entre las RATs del sistema de acuerdo al nivel de QoS requerido por cada servicio, expresado en términos de la tasa binaria R_{nom}.

A fin de poder emplear la función de utilidad u_1 para asignar un valor de utilidad a cada una de las RATs del sistema, primero es necesario realizar una estimación de la tasa binaria de transmisión que cabe esperar por parte de cada una de ellas. Para tal fin se hace necesario considerar los modos de transmisión disponibles en cada una de las RATs del sistema. Por lo general, cada RAT suele disponer de diversos modos de transmisión, cada uno de los cuales se caracteriza por el empleo de diferentes tipos

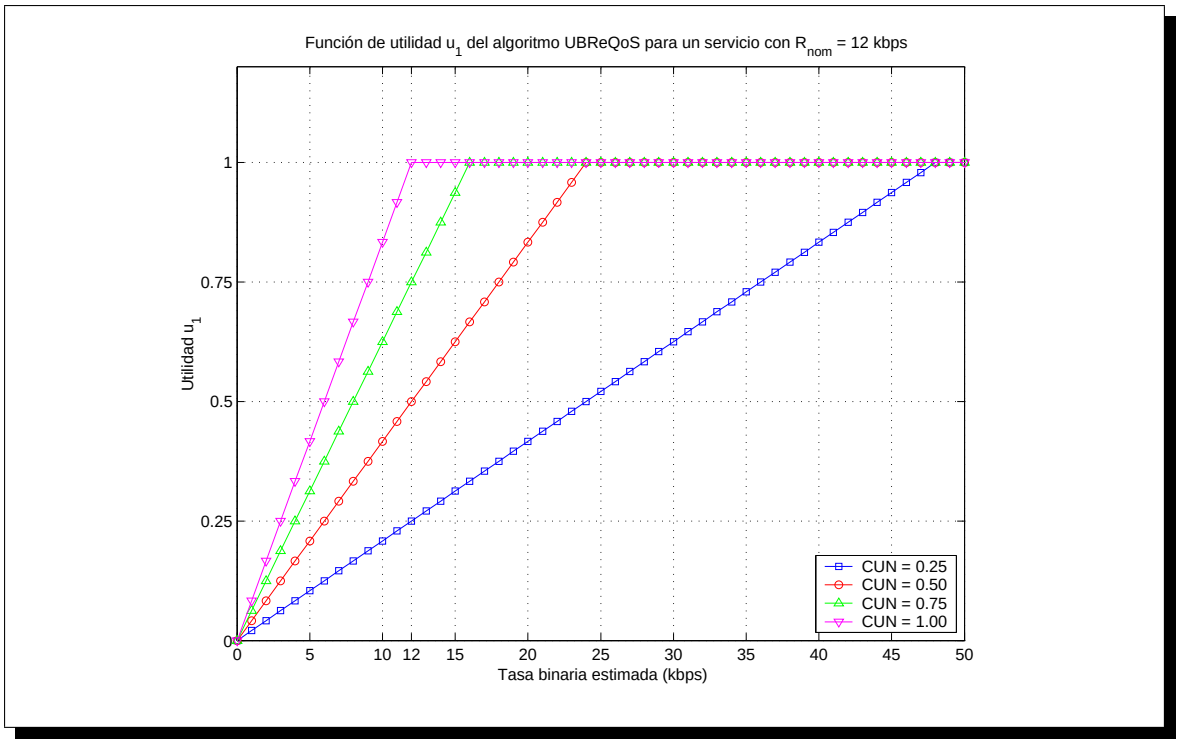


Figura 3.7: Función de utilidad u_1 del algoritmo UBReQoS para un servicio con $R_{nom} = 12$ kbps.

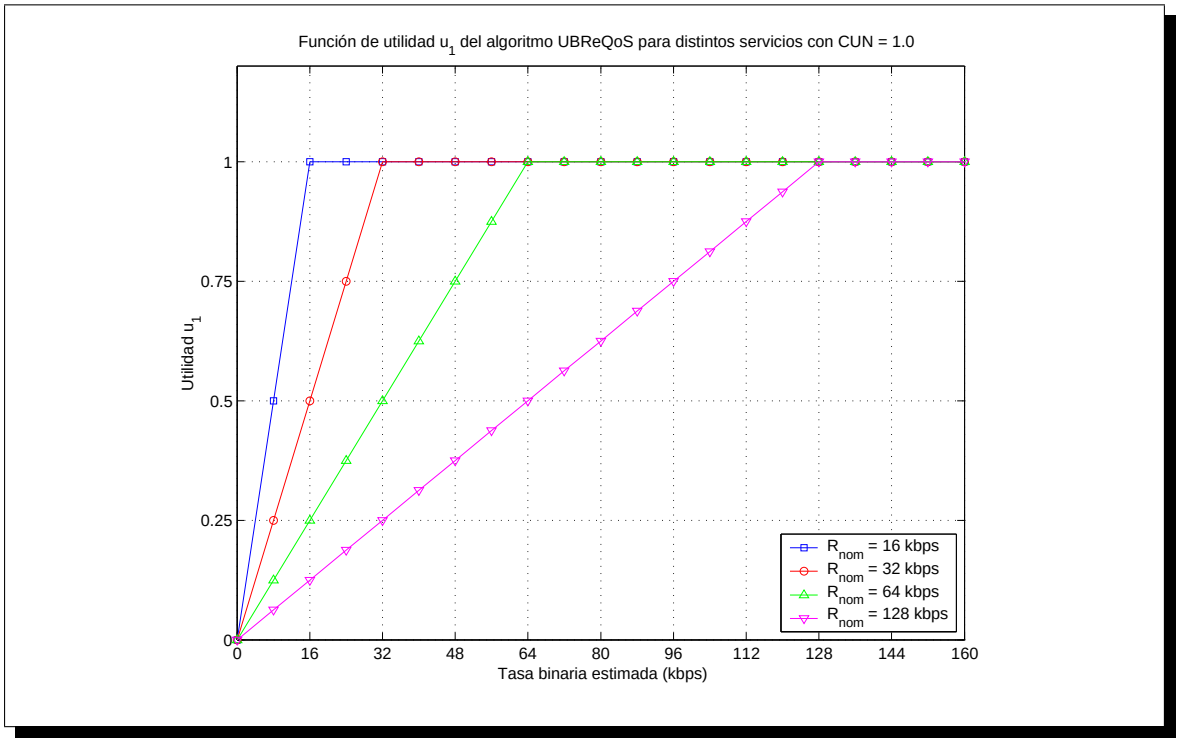


Figura 3.8: Función de utilidad u_1 del algoritmo UBReQoS para distintos servicios con CUN = 1.0.

de modulación y/o codificación de canal (ver apéndice A). Cada modo de transmisión ofrece diferente grado de compromiso entre robustez ante errores y tasa binaria de transmisión. Además, el modo de transmisión empleado suele cambiar de forma dinámica en función de la calidad experimentada en el enlace radio, de tal forma que cuando la calidad radio empeora se utilizan modos de transmisión más robustos y cuando la calidad radio es suficientemente buena se utilizan modos de transmisión menos robustos y, por lo tanto, con mayores tasas binarias de transmisión. Así pues, los diferentes modos de transmisión disponibles en una RAT ofrecen un cierto rango de velocidades de transmisión binaria. El rendimiento real de una determinada RAT vendrá, por lo tanto, fuertemente determinado por el conjunto de modos de transmisión disponibles y la frecuencia de utilización de cada uno de ellos. La forma en la cual se utiliza la función de utilidad u_1 para asignar un valor de utilidad a una RAT determinada consiste en calcular el valor de utilidad u_1 asociado a la tasa binaria de cada uno de los modos de transmisión disponibles y, a partir de dichos valores y de la probabilidad de utilización de cada modo de transmisión, calcular un valor de utilidad U_1 haciendo uso de una función de utilidad esperada de von Neumann–Morgenstern, la cual calcula la utilidad esperada en una situación de incertidumbre ponderando la utilidad de cada posible opción por su probabilidad de ocurrencia, según indica la expresión 3.5.

$$U_1 = \sum_{i=1}^{T} p(TM_i) \cdot u_1(TM_i) \tag{3.5}$$

En la expresión 3.5, T representa el número de modos de transmisión disponibles en la RAT considerada, TM_i hace referencia al i-ésimo modo de transmisión (*Tranmission Mode*, TM) de la RAT considerada, $p(TM_i)$ representa la probabilidad de que el modo de transmisión TM_i sea seleccionado, la cual puede asimilarse al porcentaje de ocasiones en las que dicho modo de transmisión ha sido seleccionado en las transmisiones previas, y $u_1(TM_i)$ representa el valor de la función de utilidad u_1 asociado a la tasa binaria del modo de transmisión TM_i. Para ilustrar el uso de dicha función, considérese un servicio de baja tasa binaria, por ejemplo, con $R_{nom} = 12$ kbps, cuya función de utilidad u_1 se calcula para CUN = 1.0 (ver figura 3.7). Para calcular la utilidad U_1 que tendría por ejemplo GPRS para dicho servicio, en primer lugar se calcularía el valor de u_1 para los modos de transmisión de GPRS. En GPRS existen cuatro modos de transmisión (desde CS-1 hasta CS-4) que ofrecen tasas binarias de 9.05, 13.4, 15.6 y 21.4 kbps (ver apéndice A). Los valores de utilidad u_1 asociados a dichas tasas binarias serían $u_1(\text{CS-1}) = 0.75$, $u_1(\text{CS-2}) = u_1(\text{CS-3}) = u_1(\text{CS-4}) = 1.0$. Suponiendo por ejemplo unas probabilidades de utilización $p(\text{CS-1}) \approx 0$ y $p(\text{CS-2}) = p(\text{CS-3}) = p(\text{CS-4}) \approx 0.33$, el valor de utilidad U_1 que tendría GPRS para el servicio considerado, haciendo uso de la expresión 3.5, sería $U_1 = 0.00 \cdot 0.75 + 0.33 \cdot 1.0 + 0.33 \cdot 1.0 + 0.33 \cdot 1.0 = 1.00$. Esto significa que, bajo las condiciones consideradas en este ejemplo, GPRS es capaz de ofrecer la máxima utilidad U_1 posible para el servicio de baja tasa binaria con $R_{nom} = 12$ kbps, por lo que el usuario de dicho servicio podría quedar satisfecho con GPRS, no habiendo por tanto, según el planteamiento propuesto por el algoritmo UBReQoS, motivo para asignar dicho usuario a otras RATs aunque éstas ofrezcan mayores prestaciones en términos de régimen binario.

Llegados a este punto, merece la pena realizar un breve inciso para comparar el planteamiento propuesto por el algoritmo UBReQoS con el sugerido al inicio de la presente sección, el cual consistía en asignar de forma fija cada servicio a una de las RATs disponibles en la red heterogénea. La idea fundamental de ambos planteamientos consiste en asignar los usuarios del sistema a aquella RAT que resulte suficiente para satisfacer sus requisitos de QoS, con el objetivo de que las RATs con mejores prestaciones estén disponibles para los usuarios de aquellos servicios que requieran niveles de QoS más elevados. Cabe, por lo tanto, preguntarse cuál es la diferencia entre ambos planteamientos. Para responder a esta pregunta considérese, una vez más, el servicio con $R_{nom} = 12$ kbps utilizado en los ejemplos anteriores. Más arriba se ha visto que dicho servicio podría en principio proveerse de forma satisfactoria a través de GPRS, ya que las tasas binarias de dicha RAT son adecuadas al valor de R_{nom} del servicio considerado. Sin embargo, supongamos que en un momento dado aumenta la demanda de dicho servicio. A medida que el número de usuarios de dicho servicio en GPRS aumenta, el nivel de interferencia global en GPRS aumentará también y, con ello, la calidad de transmisión radio se degradará, dando lugar a un aumento en el número de errores de transmisión radio. Consecuentemente, el sistema intentará compensar este deterioro de la calidad radio utilizando de forma más frecuente los modos de transmisión más robustos. Estos modos de transmisión, si bien poseen mayor protección ante errores, ofrecen velocidades binarias más reducidas. Por lo tanto, la tasas binarias de transmisión ofrecidas en GPRS experimentarán un descenso como consecuencia del aumento de la carga soportada. En el ejemplo del servicio con $R_{nom} = 12$ kbps, si GPRS experimenta cargas lo suficientemente bajas como para que la calidad de transmisión permita emplear habitualmente CS-2, CS-3 y CS-4, las tasas binarias del sistema serán superiores a 12 kbps y GPRS puede resultar suficiente para proveer dicho servicio. No obstante, si la carga aumenta lo suficiente, la calidad de transmisión radio se deteriora y comenzará a utilizarse con mayor frecuencia CS-1, cuya tasa binaria es inferior a R_{nom}. Si el uso de CS-1 aumenta lo suficiente como para que la tasa binaria esperada de GPRS quede por debajo de R_{nom}, entonces dicha RAT dejará de ser adecuada para proveer el servicio considerado, ya que en principio podría no ser capaz de garantizar el nivel de QoS requerido. Este hecho se refleja en un descenso del valor de U_1 para GPRS. En esta situación, cabe la posibilidad de que otras RATs con mayores prestaciones binarias y/o menor nivel de carga ofrezcan una tasa binaria esperada mayor y, por lo tanto, un valor más elevado de U_1, en cuyo caso los nuevos usuarios del servicio considerado pasarían de ser servidos a través de GPRS a ser servidos a través de otras RATs alternativas que ofrezcan mayores garantías de satisfacer el nivel de QoS requerido por el servicio.

Así pues, y resumiendo, es cierto que el planteamiento propuesto por el algoritmo UBReQoS y el sugerido al principio de la presente sección se basan ambos en la misma idea fundamental, consistente en asignar los usuarios del sistema a aquella RAT que resulte suficiente para satisfacer sus requisitos de QoS. No obstante, el primero es capaz de adaptar inteligentemente la asignación de usuarios de forma dinámica en función de las condiciones instantáneas de carga y QoS experimentadas en el sistema, mientras que el segundo, basado en asignaciones preestablecidas y estáticas, no podría responder ante tales situaciones y daría como resultado un deterioro del nivel de QoS.

Volviendo a la definición del algoritmo UBReQoS, la función U_1 anteriormente descrita representa la utilidad que el usuario asigna a una RAT dada del sistema en función de los requisitos de QoS (expresados en términos de tasa binaria) del servicio solicitado y de las prestaciones binarias de la RAT en cuestión. No obstante, es importante observar que el rendimiento que cabe esperar para un usuario dado en cada una de las RATs disponibles en el sistema depende, no sólo de las respectivas capacidades binarias de transmisión, sino también de la cantidad total de usuarios soportados por cada una de ellas. A medida que el número total de usuarios de una RAT aumenta, cabe esperar un deterioro del rendimiento experimentado.

Por una parte, los usuarios que se encuentran dentro de una determinada celda conectados a la misma RAT compiten entre ellos por los recursos disponibles en dicha celda. A fin de reflejar los efectos negativos que esta circunstancia produce sobre el valor de utilidad de una determinada RAT, se introduce, al igual que se hizo en la sección 3.1 para el algoritmo UBET, una función de utilidad U_2 que se define de manera idéntica,

$$U_2 = \begin{cases} 1, & N_u \leq N_c \\ N_c/N_u, & N_u > N_c \end{cases} \tag{3.6}$$

donde N_u representa el número de usuarios en la misma celda que se encuentran conectados a la misma RAT y N_c representa el número total de canales para dicha RAT en dicha celda. La finalidad y la justificación de esta función son los mismos que en el caso del algoritmo UBET (ver sección 3.1).

Por otra parte, los usuarios que producen interferencia, tanto dentro de la propia celda como fuera de ella, influyen sobre la calidad de transmisión experimentada en el enlace radio y, por lo tanto, también afectan al rendimiento final de una RAT dada. Con el objetivo de incluir de forma directa el efecto de estos usuarios sobre el valor final de utilidad, se introduce una tercera función de utilidad U_3 que se define como

$$U_3 = 1 - \sum_{i=1}^{T} p(TM_i) \cdot BLER(TM_i) \tag{3.7}$$

donde T representa el número de modos de transmisión disponibles en la RAT considerada, TM_i hace referencia al i-ésimo modo de transmisión de la RAT considerada, $p(TM_i)$ representa la probabilidad (porcentaje) de utilización del modo de transmisión TM_i, y $BLER(TM_i)$ representa el valor de BLER del modo de transmisión TM_i para la calidad de transmisión radio estimada. El valor de este último término puede obtenerse de curvas como las mostradas en las figuras A.2, A.7 y A.10 del apéndice A, las cuales representan, para cada modo de transmisión, el valor de BLER en función de la calidad radio, expresada en términos de la relación CIR/SIR. Para el cálculo del valor de CIR/SIR puede emplearse la aproximación sugerida para tal fin en la sección 3.1. Observar que a medida que aumenta el nivel de interferencia en el sistema y la calidad radio disminuye, aumenta el BLER predicho por curvas como las de las figuras A.2, A.7 y A.10, y el valor de la función de utilidad U_3 disminuye. Así pues, un elevado nivel de interferencia en una determinada RAT se refleja en un valor más reducido de U_3.

Resumiendo, el modo de operación del algoritmo UBReQoS es el siguiente. Cada vez que es necesario decidir a qué RAT debe asignarse un usuario, se calcula, para cada una de las RATs disponibles en la posición geográfica del usuario, el valor de las funciones de utilidad anteriormente descritas. Para ello, se calcula en primer lugar la función de utilidad u_1 a partir del valor del parámetro R_{nom} asociado al servicio solicitado por el usuario y del valor del parámetro CUN (que puede considerarse como un parámetro de diseño del algoritmo). Haciendo uso de la función u_1 se calcula el valor de utilidad correspondiente a las tasas binarias de los modos de transmisión disponibles en cada una de las RATs del sistema. El valor de utilidad U_1 de una determinada RAT se obtiene a partir de los valores de utilidad u_1 correspondientes a sus respectivos modos de transmisión, tal y como se indicó en la expresión 3.5. Para calcular el valor de la función de utilidad U_2 de una RAT dada, simplemente es necesario contar el número de usuarios que se hayan conectados a dicha RAT en la misma celda en la cual se encuentra el usuario para el cual se va a decidir la RAT a asignar, y aplicar la expresión 3.6. Por último, para obtener el valor de la función de utilidad U_3 de una RAT determinada, puede estimarse el nivel de interferencia experimentado en dicha RAT y obtener así un valor para la relación CIR/SIR, o bien puede obtenerse el número de usuarios interferentes intercelda e intracelda (si procede) y, consultando una tabla almacenada, determinar el valor de CIR/SIR correspondiente. En cualquier caso, el valor de CIR/SIR obtenido se transforma en un valor de BLER para cada uno de los modos de transmisión de la RAT considerada y, aplicando la expresión 3.7, se obtiene un valor de U_3. Finalmente, el valor de utilidad total U asociado a una RAT se obtiene como el producto de las tres funciones de utilidad anteriores, es decir, $U = U_1 \cdot U_2 \cdot U_3$. Repitiendo este procedimiento para todas las RATs disponibles en la posición geográfica del usuario, se obtiene un valor de utilidad U para cada una de ellas. El usuario en cuestión será asignado a aquella RAT que ofrezca el valor más elevado de U. En caso de que varias de las RATs disponibles en el sistema ofrezcan el máximo valor de utilidad posible, el usuario será asignado a aquella cuyas capacidades de transmisión sean menores. Esta manera de proceder tiene como finalidad que los recursos de las RATs con mayores prestaciones estén disponibles para los usuarios de aquellos servicios para los cuales el máximo valor de utilidad sólo pueda ser ofrecido por dichas RATs.

Antes de concluir la presente sección, y dada la importancia del parámetro R_{nom} sobre el funcionamiento del algoritmo UBReQoS, merece la pena explicar detalladamente el significado de dicho parámetro así como la forma en la cual ha sido determinado su valor en este trabajo para diferentes tipos de servicios. Para ello se ha realizado una distinción entre servicios en tiempo real y servicios en tiempo no real. En el caso de servicios en tiempo real, éstos se caracterizan, a grandes rasgos, por la necesidad de transmitir paquetes de datos de un determinado tamaño dentro de un intervalo de tiempo prefijado para cada paquete. El cociente entre el tamaño del paquete de datos y el tiempo disponible para su transmisión establece la tasa binaria mínima requerida para que dicho paquete pueda transmitirse dentro del plazo de tiempo establecido. Promediando las tasas binarias requeridas por un número suficientemente amplio de paquetes, se obtiene un valor representativo de la tasa binaria media requerida por dicho servicio, la cual puede emplearse como valor del parámetro R_{nom} para la aplicación

del algoritmo UBReQoS. En el caso de servicios en tiempo no real, la identificación del valor de R_{nom} no es algo obvio ni inmediato. Por ello, a continuación se proponen varias alternativas para determinar el valor de R_{nom} para servicios en tiempo no real. La idea fundamental en la que se basan consiste en determinar un valor de tamaño de paquete que sea representativo de los paquetes de datos generados por el servicio considerado, y dividir dicho tamaño entre el intervalo de tiempo de transmisión que el usuario considera tolerable. Para ilustrar los métodos propuestos se consideran dos servicios en tiempo no real diferentes, concretamente, navegación web y correo electrónico. El motivo por el cual se han seleccionado estos dos servicios es porque son los servicios en tiempo no real que se implementan en la plataforma de simulación empleada en este trabajo para evaluar el rendimiento de los algoritmos propuestos.

En primer lugar, se hace necesario disponer de una secuencia de paquetes de datos generados por el tipo de servicio bajo estudio sobre la cual realizar los análisis oportunos a fin de obtener un valor de R_{nom} para dicho servicio. En una aplicación real dicha secuencia de paquetes podría obtenerse capturando varias sesiones de tráfico en un sistema real y utilizando el tamaño de los paquetes capturados para el análisis. En este trabajo se ha utilizado una secuencia de paquetes generados con los modelos de tráfico de navegación web y correo electrónico implementados en la plataforma de simulación empleada, los cuales se describen en el capítulo 4. Para tal fin se ha generado una secuencia de $5 \cdot 10^6$ paquetes de datos (por paquete de datos se entiende una página web o un correo electrónico). Las tablas 3.1 y 3.2 muestran las estadísticas de las secuencias generadas. Podría tomarse por ejemplo el valor medio como valor representativo del tamaño de los paquetes generados. No obstante, cuando en una secuencia existen valores con magnitudes considerablemente dispares, el valor medio puede resultar estadísticamente poco representativo ya que unos pocos valores muy pequeños o muy grandes pueden llegar a modificar de forma relativamente significativa el valor medio de la secuencia. Este es precisamente el caso de las secuencias generadas, tal y como muestra la gran diferencia observada entre los valores mínimo y máximo de las tablas 3.1 y 3.2. En esta circunstancia, el valor mediano suele ser un valor más representativo de los valores habituales en la secuencia. Tal y como puede apreciarse en las tablas 3.1 y 3.2, el valor mediano está más próximo al valor más frecuente de la secuencia que el valor medio. Así pues, consideraremos el valor mediano de las secuencias generadas como valor representativo de los tamaños de los paquetes de datos.

Adicionalmente, se ha calculado la función de distribución acumulativa (*Cumulative Distribution Function*, CDF) de los tamaños de los paquetes generados. Las figuras 3.9 y 3.10 muestran, respectivamente, dichas CDFs para los paquetes generados por los modelos de tráfico de navegación web y correo electrónico empleados. Es importante observar que, por definición, el valor mediano de una secuencia coincide con el punto correspondiente a una probabilidad del 50 % en la CDF de dicha secuencia. Es decir, el valor mediano de una secuencia representa, por definición, el valor máximo observado en el 50 % de las muestras. Las figuras 3.9 y 3.10 ilustran claramente este hecho. Aunque para la evaluación del algoritmo UBReQoS se ha seleccionado el tamaño de paquete correspondiente al valor mediano, es decir, al 50 % de la CDF, también se analizará el efecto de considerar otros valores, concretamente, los correspondientes al 70 % y 90 %.

Valor mínimo	8000 bits
Valor máximo	160802112 bits
Valor medio	115633 bits
Valor más frecuente	8104 bits
Valor mediano	20184 bits
Valor 50 % CDF	20184 bits
Valor 70 % CDF	37856 bits
Valor 90 % CDF	124400 bits

Tabla 3.1: Estadísticas de los tamaños de página web.

Valor mínimo	24 bits
Valor máximo	83675928 bits
Valor medio	285955 bits
Valor más frecuente	35080 bits
Valor mediano	47448 bits
Valor 50 % CDF	47448 bits
Valor 70 % CDF	107744 bits
Valor 90 % CDF	618832 bits

Tabla 3.2: Estadísticas de los tamaños de correo electrónico.

Una vez determinado, por cualquiera de los procedimientos anteriores, un valor representativo del tamaño de los paquetes de datos, el valor de R_{nom} se obtiene dividiendo dicho valor entre un determinado intervalo de tiempo. Siguiendo las indicaciones dadas en [154], se establece como requisito subjetivo de QoS para navegación web y correo electrónico que tanto las páginas web como los correos electrónicos sean transmitidos en un tiempo máximo de 4 segundos. Considerando dicho intervalo de tiempo, el valor del parámetro R_{nom} que se obtiene a partir del tamaño obtenido para el 50 %, 70 % y 90 % de las CDFs es, respectivamente, de 5 kbps, 9.5 kbps y 31 kbps para navegación web y de 12 kbps, 27 kbps y 155 kbps para correo electrónico. Los valores obtenidos para correo electrónico se deben al elevado tamaño de los correos con archivos adjuntos, posibilidad que ha sido modelada en el simulador empleado (ver capítulo 4).

Los valores anteriores obtenidos mediante el procedimiento propuesto no deberían ser interpretados en sentido estricto como las tasas binarias realmente requeridas por una aplicación real de navegación web o de correo electrónico. Hay que tener en cuenta que dichos valores han sido obtenidos considerando unos modelos concretos y podrían variar de manera significativa si se consideraran otros modelos distintos o si se obtuvieran a partir sesiones de tráfico capturadas de una red real. Además, las tasas binarias requeridas por los servicios considerados podrían variar en función de condiciones específicas o de los objetivos de QoS propuestos. Por ejemplo, un usuario podría considerar satisfactoria la provisión de un servicio de navegación web si las páginas se transmiten en un plazo de tiempo inferior a los 4 segundos señalados en [154], pero otro usuario al cual se le tarificara el mismo servicio a un precio mayor podría considerarlo satisfactorio sólo si dicho intervalo de tiempo fuera por ejemplo inferior a 1 segundo, con lo cual la tasa binaria del mismo servicio sería mayor en el caso de este último usuario. Este tipo de ambigüedades no tendrían lugar en el caso de servicios en tiempo real, los cuales tienen asociada una tasa binaria claramente definida que determina si los paquetes de datos son recibidos a tiempo y, por lo tanto, si la provisión del servicio puede objetivamente considerarse satisfactoria. Así pues, es importante tener en cuenta que los valores de R_{nom} anteriormente obtenidos por el procedimiento descrito corresponden a unas condiciones de aplicación particulares (las consideradas en este trabajo) que podrían variar de forma significativa en función del contexto considerado.

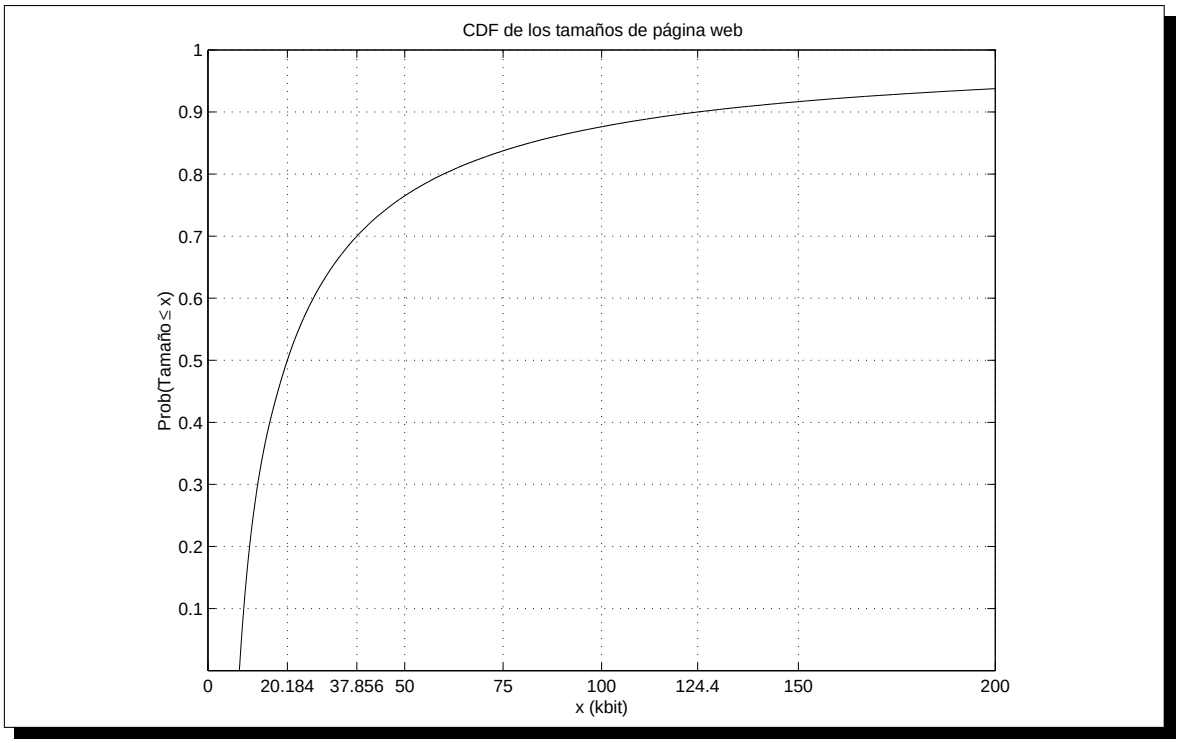


Figura 3.9: CDF de los tamaños de página web.

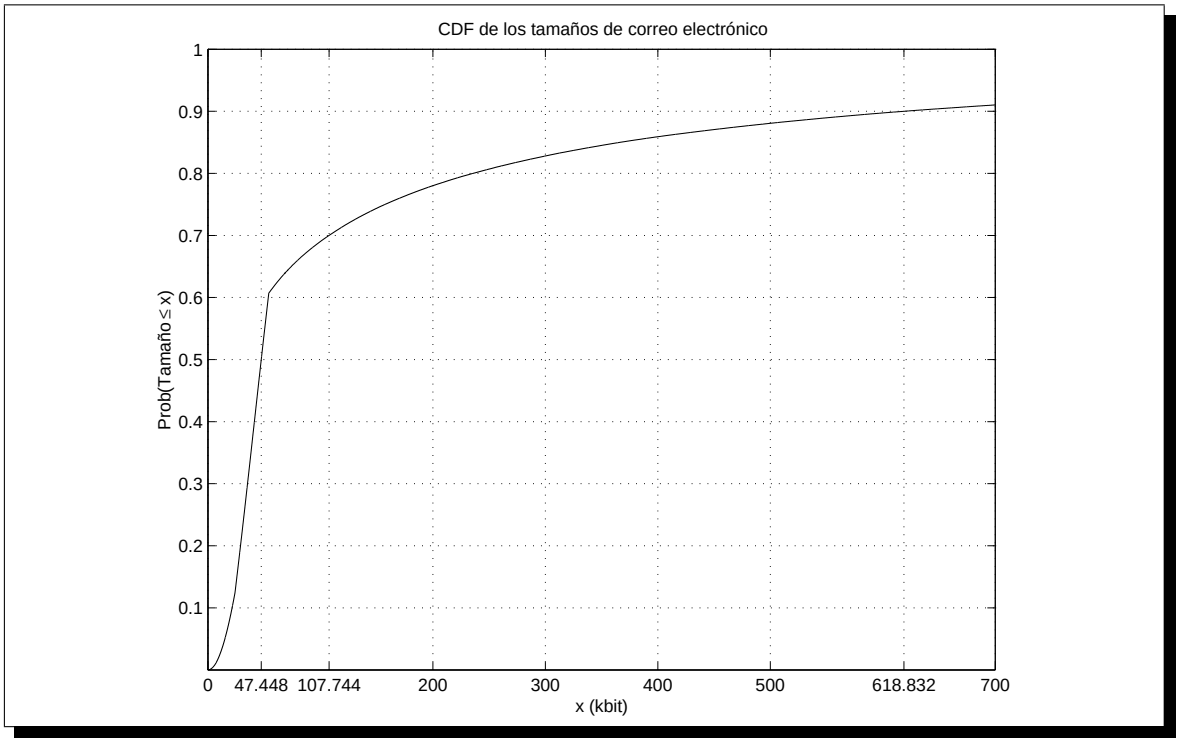
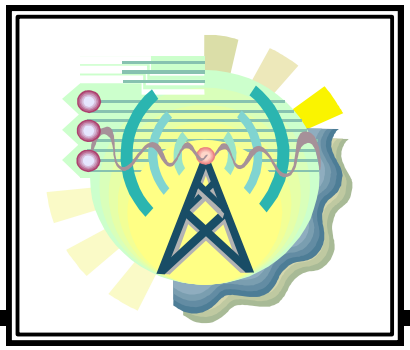


Figura 3.10: CDF de los tamaños de correo electrónico.

3.3. Sumario

El objetivo del presente trabajo consiste en el desarrollo de nuevas técnicas CRRM que permitan distribuir eficientemente el tráfico de una red radio heterogénea entre las diversas alternativas de acceso radio disponibles, garantizando los niveles de QoS requeridos al mismo tiempo que se optimiza el uso de los recursos radio del sistema. En este capítulo se han propuesto dos nuevos algoritmos diseñados para tal fin. El primero de ellos pretende obtener elevados niveles de QoS asignando los usuarios del sistema a aquella RAT que, en cada momento, ofrece el valor de *throughput* estimado más elevado. Dicha estimación se realiza a partir de la calidad de transmisión experimentada en el enlace radio y de la disponibilidad de recursos. El segundo de los algoritmos propuestos se diseña para distribuir el tráfico de la red radio heterogénea entre las RATs del sistema en función de las prestaciones de cada una de ellas y del nivel de QoS requerido por el usuario. Los usuarios del sistema son asignados a aquellas RATs que resultan suficientes para satisfacer sus requisitos de QoS, con la intención de que las RATs con mejores prestaciones estén disponibles para los usuarios de aquellos servicios que requieran niveles de QoS más elevados. El objetivo de este último algoritmo es, así pues, el de ofrecer niveles de QoS adecuados para todos los usuarios del sistema.



CAPÍTULO 4

Modelos y Herramientas de Simulación

Existen diversas alternativas para evaluar el rendimiento de una determinada técnica en un sistema de comunicaciones móviles. Una primera opción consistiría en implementarla sobre un sistema real y realizar posteriormente medidas sobre dicho sistema. Con esta opción todos los factores intervienen sobre los resultados finales, permitiendo una evaluación objetiva y fiable del rendimiento. Sin embargo, esta opción es compleja y supone un elevado coste económico. Una segunda opción consiste en diseñar y emplear modelos analíticos. La ventaja de los modelos analíticos reside en el hecho de que es posible establecer, mediante fórmulas matemáticas, una relación clara y directa entre los diversos parámetros que determinan el comportamiento de un sistema así como su grado de relevancia sobre el rendimiento final. Sin embargo, la obtención de una solución cerrada puede llegar a resultar extraordinariamente compleja y a menudo es inevitable realizar bastantes suposiciones y simplificaciones para que el problema sea analíticamente tratable. Este hecho limita la fiabilidad y exactitud del modelo analítico. La elevada y creciente complejidad de los sistemas de comunicaciones móviles aumenta la dificultad de estudiar nuevas técnicas mediante modelos analíticos. Una tercera alternativa consiste en la simulación por ordenador. Para ello, los elementos y el funcionamiento del sistema se modelan mediante *software* que se ejecuta en una computadora. En general, este método necesita un elevado número de ejecuciones para que los resultados obtenidos sean estadísticamente representativos del comportamiento real del sistema, lo cual requiere un elevado tiempo de simulación. La cantidad de suposiciones y aproximaciones necesarias depende de la complejidad del modelo de simulación, pero suele ser notablemente menor que en el caso del estudio mediante modelos analíticos, lo cual garantiza una mayor exactitud y fiabilidad de los resultados obtenidos. Esta alternativa supone un buen compromiso entre coste, tiempo, exactitud y complejidad.

La evaluación del rendimiento de los algoritmos estudiados en este trabajo se lleva a cabo mediante simulación por ordenador. El presente capítulo estará dedicado, así pues, a la presentación y descripción de los modelos empleados así como de las herramientas utilizadas para llevar a cabo las simulaciones realizadas en este trabajo.

4.1. Simulación a nivel de enlace y de sistema

El estudio por simulación de un sistema de comunicaciones móviles suele desglosarse en dos niveles con diferente resolución temporal: *nivel de enlace* y *nivel de sistema*. A nivel de enlace se modela el comportamiento del enlace radio bit a bit, considerando aspectos tales como la codificación de canal, la modulación o los efectos de la radiopropagación en una escala de tiempos relativamente reducida (desvanecimiento rápido). A nivel de sistema se modela el comportamiento de la red celular incluyendo factores como, por ejemplo, la movilidad de los usuarios, sus relaciones de interferencia o los efectos producidos por el canal de transmisión en una escala de tiempos relativamente amplia (atenuación y desvanecimiento lento). Así pues, a nivel de sistema se consideran aspectos que varían en una escala de tiempos más o menos amplia mientras que a nivel de enlace la escala temporal es considerablemente más reducida.

Los simuladores a nivel de enlace y a nivel de sistema se ejecutan de forma totalmente independiente. La interacción entre ambos niveles de simulación se consigue mediante el empleo de unas interfaces denominadas *Look-Up Tables* (LUTs). Las LUTs son un conjunto de tablas, como las mostradas en las figuras A.2, A.7 y A.10 del apéndice A, que representan de forma simplificada los resultados de la simulación a nivel de enlace, relacionando la calidad experimentada en el enlace radio en términos, por ejemplo, de la relación portadora a interferencia (*Carrier-to-Interference Ratio*, CIR) o señal a interferencia (*Signal-to-Interference Ratio*, SIR), con un determinado parámetro de calidad como, por ejemplo, la tasa de bloques erróneos (*BLock Error Rate*, BLER). Este conjunto de tablas sirve de información de entrada para las simulaciones a nivel de sistema, durante las cuales serán consultadas para obtener la información deseada relativa al nivel de enlace. Este procedimiento representa una forma sencilla, pero eficiente a la vez, de incluir en las simulaciones a nivel de sistema, basadas en una escala de tiempos relativamente amplia, aquellos efectos introducidos por el canal de transmisión en una escala de tiempos mucho más reducida.

El principal motivo para desglosar las simulaciones en estos dos niveles es la eficiencia en términos de coste computacional y tiempo requerido. Los aspectos relativos al enlace radio como, por ejemplo, la codificación de canal o la modulación, deben simularse con un elevado nivel de resolución temporal (bit a bit) y, por lo tanto, llevan asociada una importante carga computacional. Agrupando estos aspectos en la simulación a nivel de enlace y representando sus efectos de manera simplificada mediante un conjunto de tablas (LUTs) que posteriormente son consultadas durante la simulación a nivel de sistema (con un nivel de resolución temporal considerablemente menor) se reduce la cantidad de recursos y de tiempo requeridos para efectuar la simulación.

Para las simulaciones llevadas a cabo en este trabajo se han empleado LUTs extraídas de diferentes fuentes bibliográficas (ver apéndice A), las cuales han sido integradas en una sofisticada plataforma de simulación a nivel de sistema que ha sido desarrollada para los fines de este trabajo. Dicha plataforma de simulación, bautizada con el nombre de SPHERE (*Simulation Platform for HEterogeneous wiREless systems*) [155, 156], integra tres avanzados simuladores a nivel de sistema de las tecnologías de acceso radio GPRS, EDGE y HSDPA (*High Speed Downlink Packet Access*), esta última evolución del actual estándar UMTS generalmente calificada como tecnología 3.5G (en el apéndice A se ofrece un resumen de las características más relevantes de estas tres tecnologías). Estos tres simuladores implementan fielmente las especificaciones de la interfaz radio de sus respectivas normas, permitiendo una evaluación precisa tanto del rendimiento de las técnicas CRRM evaluadas en este trabajo como de la calidad final que percibirán los usuarios en una red heterogénea de comunicaciones móviles.

4.2. Plataforma de simulación

La plataforma de simulación se ha implementado en lenguaje C++ y haciendo uso de la librería CNCL (*Communication Networks Class Library*) [157]. CNCL es una librería de clases para simulación por eventos discretos escrita en C++ y pensada para proveer funciones de simulación comunes a cualquier red de comunicaciones. Incluye funcionalidades para la generación de números aleatorios, evaluación estadística, creación de estructuras contenedoras como listas, *arrays*, colas o pilas y la simulación controlada por eventos. El elemento principal de esta librería es el *event scheduler*, entidad central de control encargada de ordenar los eventos en función de su instante de ejecución y su prioridad así como de entregarlos, en el instante de simulación correspondiente, a los manejadores de eventos responsables de su gestión.

La figura 4.1 muestra la estructura lógica de la plataforma de simulación, dividida en cuatro entidades. En las secciones siguientes se describe de forma detallada todos los aspectos vinculados a cada una de las cuatro entidades mostradas en dicha figura.

4.2.1. Entorno celular

La plataforma de simulación se basa en un entorno celular formado por 27 celdas omnidireccionales con un radio de 500 m y distribuidas según una estructura de clúster formada por 3 celdas. Las estaciones base se consideran ubicadas en el centro de las celdas y en cada una de ellas existe cobertura de todas las tecnologías de acceso radio disponibles en el sistema. Para evitar los efectos de los bordes en las celdas exteriores se ha aplicado la técnica *wrap-arround*, consistente en *plegar* imaginariamente el plano que contiene las celdas del sistema hasta hacer coincidir el borde superior con el inferior y el borde izquierdo con el derecho. El resultado es equivalente a repetir la estructura de 27 celdas alrededor de sí misma, tal y como se muestra en la figura 4.2.

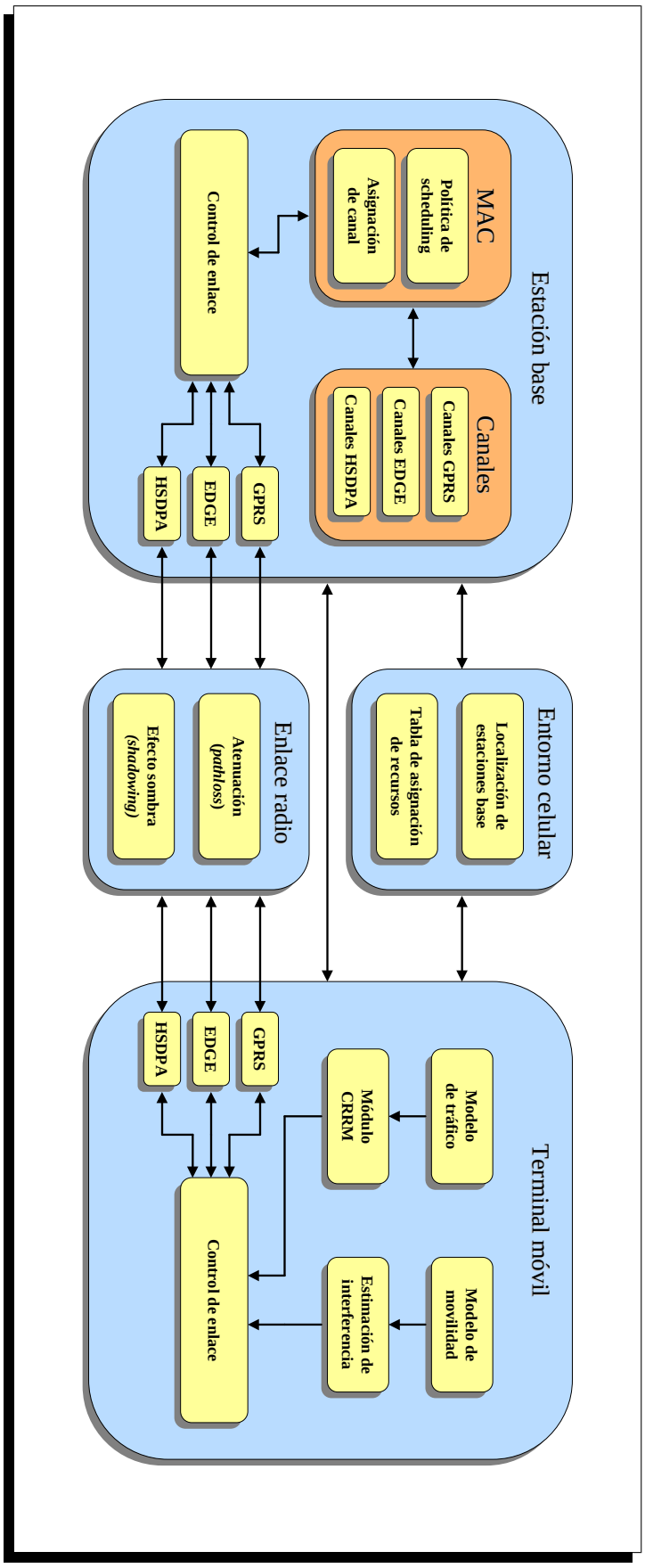


Figura 4.1: Estructura lógica de la plataforma de simulación.

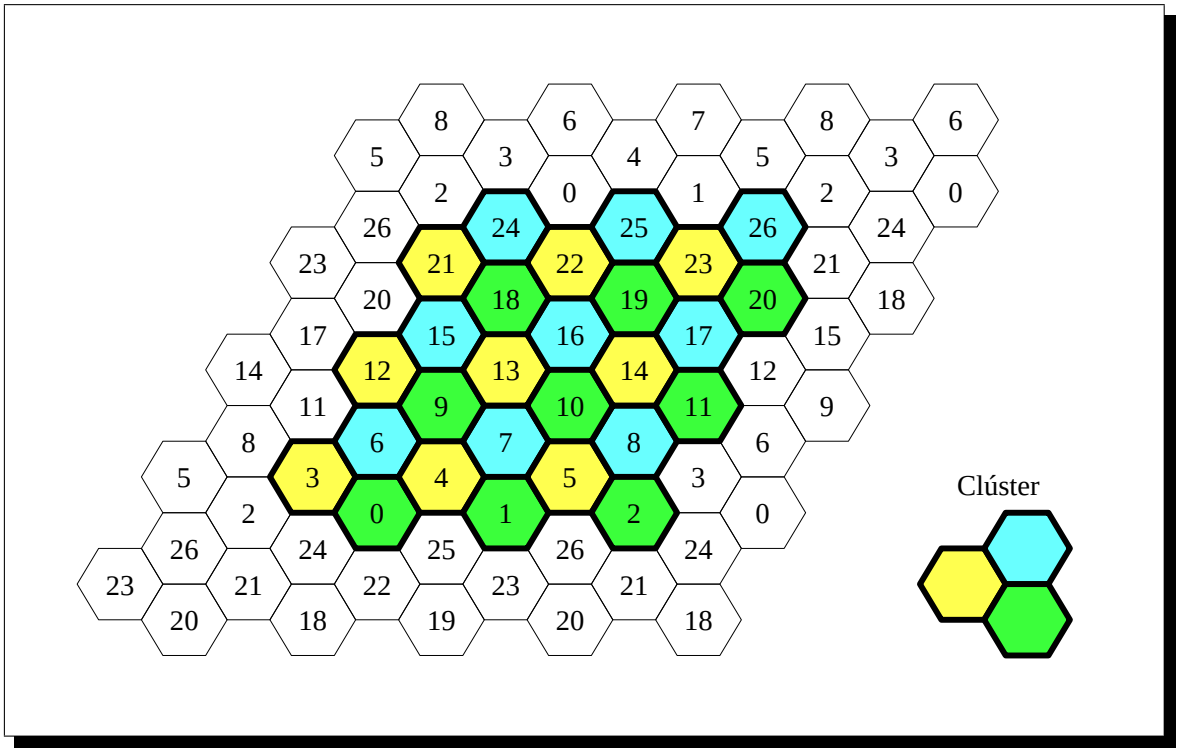


Figura 4.2: Estructura celular de la plataforma de simulación.

En la entidad *Entorno celular* se almacena información relativa a la posición física de las estaciones base así como una lista actualizada de los recursos asignados en cada una de ellas. Esta información, almacenada en forma de tablas, se utiliza para la estimación de interferencias, lo cual requiere conocer qué canales están activos en cada estación base interferente y la distancia entre cada una de ellas y el receptor interferido.

4.2.2. Enlace radio

En la entidad *Enlace radio* se modelan los fenómenos de propagación radio entre los extremos transmisor y receptor, caracterizando las variaciones producidas sobre el nivel de señal en una escala de tiempos relativamente amplia, es decir, la atenuación o *pathloss* y el denominado efecto sombra o *shadowing*.

Para el cálculo de atenuación se emplea el modelo de propagación Okumura-Hata en su extensión COST 231. El modelo inicial se basaba en una campaña de medidas realizadas por Y. Okumura [158] en la ciudad de Tokyo a partir de las cuales M. Hata [159] elaboró un modelo matemático ajustado a dichas medidas. Posteriormente, dada la limitación en frecuencia del modelo original, la acción COST 231 [160] introdujo modificaciones para adecuarlo a las frecuencias de operación de los sistemas 3G. Según este modelo de propagación, la atenuación L en dB viene dada por la expresión

$$\begin{aligned} L &= A + B \cdot \log_{10}(f_c) - 13{,}82 \cdot \log_{10}(h_b) \\ &\quad -a(h_m) + [C - 6{,}55 \cdot \log_{10}(h_b)] \cdot \log_{10}(d) \end{aligned} \tag{4.1}$$

donde A, B y C son constantes, f_c es la frecuencia portadora en MHz, h_b es la altura de la estación base en metros, h_m es la altura del terminal móvil en metros, $a(h_m)$ es la función de ganancia de la antena del terminal móvil y d es la distancia entre estación base y terminal móvil en kilómetros.

El valor de los parámetros A y B se establece según la tabla 4.1. Estos valores han sido determinados a partir de medidas experimentales. El valor del parámetro C, que representa la dependencia del modelo con la distancia, generalmente está comprendido entre 44 y 47. El valor por defecto utilizado más a menudo es $C = 44{,}9$.

f_c	150-1500 MHz	1500-2000 MHz
A	69.55	46.3
B	26.16	33.9

Tabla 4.1: Valor de los parámetros A y B para el modelo Okumura-Hata COST 231.

La expresión de la función $a(h_m)$ depende del tipo de ciudad. Para pequeñas y medianas ciudades

$$a(h_m) = [1{,}1 \cdot \log_{10}(f_c) - 0{,}7] \cdot h_m - [1{,}56 \cdot \log_{10}(f_c) - 0{,}8] \tag{4.2}$$

mientras que para grandes ciudades

$$a(h_m) = \begin{cases} 8{,}29 \cdot [\log_{10}(1{,}54 \cdot h_m)]^2 - 1{,}1, & f_c \leq 200MHz \\ 3{,}2 \cdot [\log_{10}(11{,}75 \cdot h_m)]^2 - 4{,}97, & f_c \geq 400MHz \end{cases} \tag{4.3}$$

Es interesante observar que normalmente estas funciones no tienen un peso significativo sobre el valor de atenuación debido a que la altura de la antena del terminal móvil generalmente ronda, de manera aproximada, los 1.5 metros. Para $h_m = 1{,}5$ el valor devuelto por estas funciones es prácticamente cero y muy poco sensible a pequeñas variaciones en la altura de la antena del terminal móvil.

El modelo es válido en los siguientes intervalos:

- Frecuencia portadora: 150 MHz $\leq f_c \leq$ 2000 MHz.
- Altura de la estación base: 30 m $\leq h_b \leq$ 200 m.
- Altura del terminal móvil: 1 m $\leq h_m \leq$ 10 m.
- Distancia entre estación base y terminal móvil: 1 km $\leq d \leq$ 20 km.

Aunque el modelo está basado en medidas realizadas sobre distancias superiores a 1 km, éste sigue siendo válido para distancias inferiores [161].

El concepto de atenuación representa un valor medio para las pérdidas experimentadas sobre una determinada distancia. Sin embargo, para una misma distancia entre transmisor y receptor puede obtenerse diversos valores instantáneos de pérdida. Este fenómeno se incluye mediante el denominado efecto sombra, el cual representa pérdidas adicionales por la presencia de obstáculos entre transmisor y receptor. Las pérdidas introducidas por el efecto sombra pueden modelarse mediante un proceso aleatorio con una distribución normal de valor medio 0 dB y una desviación estándar entre 4 y 12 dB dependiendo del entorno de propagación (se ha considerado un valor de 6 dB). El efecto sombra es un proceso espacialmente correlado, es decir, el valor de las pérdidas debidas al efecto sombra que experimenta un terminal móvil en una posición concreta está correlado con el experimentado en una posición cercana. El modelo utilizado para la correlación del efecto sombra con la distancia es el descrito en [162], según el cual la correlación entre dos posiciones separadas por una distancia x viene dada por

$$R(x) = \sigma^2_{decorr} \cdot \exp\left(-\frac{|x|}{x_0}\right) \tag{4.4}$$

donde σ_{decorr} es la desviación estándar y $x_0 = d_{decorr}/\ln(2)$, siendo d_{decorr} la distancia de decorrelación, la cual se define como la distancia requerida para obtener un cierto grado de decorrelación entre dos muestras del proceso asociado al efecto sombra. Un valor generalmente considerado para definir la distancia de decorrelación es 0,5, lo cual significa que d_{decorr} se define como aquel valor para el cual la correlación normalizada es igual a 0,5. La distancia de decorrelación depende de las condiciones del entorno. Para un entorno macrocelular, un valor apropiado es $d_{decorr} = 20\ m$ [163].

4.2.3. Estación base

En la plataforma de simulación se modela la transmisión en el enlace descendente, es decir, en sentido estación base a terminal móvil, motivo por el cual la entidad *Estación base* es responsable de las funciones de control de acceso al medio (*Medium Access Control*, MAC), concretamente de la política de *scheduling* aplicada y del mecanismo de asignación de canales radio.

Cuando se produce una solicitud de canal de transmisión para una tecnología de acceso radio determinada se examina la lista de canales disponibles en la misma. Si los hay, se selecciona uno de ellos y se produce la asignación. Para la asignación de canal se ha implementado un algoritmo de selección aleatoria mediante el cual se escoge al azar un canal del conjunto de canales disponibles. Si no hay ningún canal disponible, el terminal móvil pasa a una cola de espera hasta que alguno de los usuarios del sistema finaliza su transmisión y libera el canal, momento en el cual dicho canal se asigna a uno de los usuarios de la cola de espera. Para los usuarios en las colas de GPRS y EDGE se aplica una política de *scheduling* de tipo FCFS (*First-Come First-Served*), en la cual las solicitudes de canal se satisfacen en el mismo orden en que surgen. Los usuarios de la cola de HSDPA acceden a los canales de transmisión según una política de tipo *Round Robin* con un cuanto temporal igual a la duración de un intervalo de transmisión o TTI (2 ms). El elemento *Control de enlace* se describirá más adelante.

4.2.4. Terminal móvil

4.2.4.1. Modelo de movilidad

La posición inicial de cada terminal móvil dentro de una celda se establece de forma aleatoria de acuerdo a una distribución uniforme y el desplazamiento dentro de dicha celda se realiza a velocidad constante. Para integrar la movilidad de los terminales móviles (carácter continuo) con la simulación por eventos (carácter discreto), el movimiento de un terminal móvil se modela como una sucesión de posiciones aleatorias en instantes discretos de tiempo. La separación entre dos posiciones consecutivas es constante e igual a la distancia de decorrelación empleada en el modelo de *shadowing*. Las ubicaciones intermedias entre dos posiciones consecutivas se obtienen por interpolación lineal. En cada una de estas posiciones, la dirección del movimiento se decide añadiendo cierto ángulo aleatorio a la dirección previa, el cual se obtiene a partir de una distribución normal con media igual a cero y con una varianza que depende de la velocidad del terminal móvil. Este modelo de movilidad da lugar a una densidad de usuarios uniforme en la celda y ha demostrado ser bastante realista al ser contrastado con datos de un sistema real [164]. La figura 4.3 muestra un ejemplo del movimiento de un terminal móvil dentro de una celda.

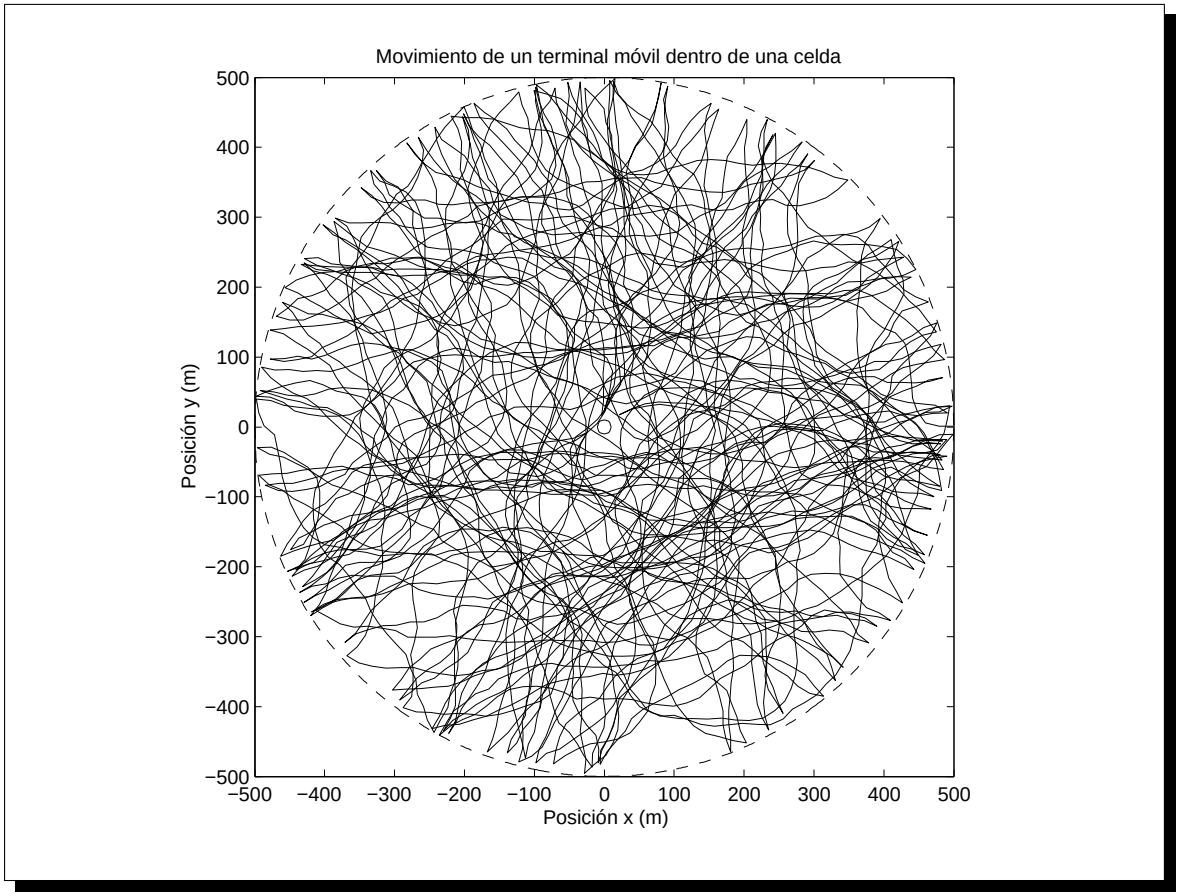


Figura 4.3: Movimiento de un terminal móvil dentro de una celda.

4.2.4.2. Estimación de interferencia

En la plataforma de simulación, cada terminal móvil es responsable de estimar el nivel de interferencia recibida para obtener la calidad de su propio canal de transmisión, expresada en términos de relación CIR o SIR. Para GPRS y EDGE se ha empleado el acrónimo CIR, ya que en dichos sistemas, basados en un acceso múltiple de tipo TDMA/FDMA, el nivel de señal útil se identifica con el nivel de señal portadora recibido. Para HSDPA, basado en un acceso múltiple de tipo CDMA, se ha preferido emplear el acrónimo SIR, ya que en este tipo de sistemas el nivel de señal útil corresponde a una parte de la potencia de portadora recibida. El resto de potencia corresponde a las señales de otros usuarios o a canales de control, señalización, etc.

En ambos casos el nivel de señal útil recibida por un terminal móvil viene dado por la potencia de transmisión de la señal en su estación base y por la distancia que separa a ambos. Dicha distancia vendrá determinada por la ubicación física tanto de la estación base como del terminal móvil.

En GPRS y en EDGE (sistemas FDMA/TDMA), el nivel de potencia interferente se calcula de manera similar pero considerando la potencia de transmisión de cada una de las estaciones base cocanal interferentes, sus canales interferentes que están activos y la distancia entre el terminal móvil interferido y cada una de estas estaciones base interferentes. Una vez obtenido el valor de señal recibido de cada estación base interferente se suman todos ellos para obtener finalmente el valor total de potencia interferente. Así pues, para GPRS y para EDGE

$$CIR_{GPRS/EDGE} = \frac{\frac{P_i}{L_P^{ii} \cdot L_S^{ii}}}{\sum_{j \in \Omega} \frac{P_j}{L_P^{ij} \cdot L_S^{ij}} + N_0 \cdot W} \tag{4.5}$$

donde P_i es la potencia de transmisión de la señal deseada en la celda de referencia (celda i), L_P^{ii} y L_S^{ii} son respectivamente las pérdidas por *pathloss* y *shadowing* del enlace radio entre transmisor y receptor en la celda de referencia, Ω es el conjunto de canales interferentes que están activos en las celdas interferentes, P_j es la potencia de transmisión en las celdas interferentes, L_P^{ij} y L_S^{ij} son las pérdidas por *pathloss* y *shadowing* de los enlaces entre los transmisores interferentes en las celdas j y el receptor interferido en la celda i y $N_0 \cdot W$ representa el ruido térmico, siendo N_0 la densidad espectral de ruido y W el ancho de banda de transmisión del canal.

En sistemas CDMA como HSDPA, los códigos de canalización (ver apéndice A) de los usuarios de una misma celda son perfectamente ortogonales entre sí, pero debido al efecto multicamino esta ortogonalidad se reduce [165], ocasionando cierta componente adicional de interferencia intracelda que se modela mediante un factor de ortogonalidad α [153]. En ausencia de efecto multicamino, $\alpha = 1$. Si se dan dos componentes multicamino de igual potencia, $\alpha = 0{,}5$. En el peor de los casos, $\alpha = 0$ (se destruye por completo la ortogonalidad). Típicamente, α toma valores entre 0.4 y 0.9 [166]. Así pues,

para HSDPA

$$SIR_{HSDPA} = \frac{\dfrac{P_i}{L_P^{ii} \cdot L_S^{ii}}}{\displaystyle\sum_{i' \in \Omega'} \frac{P_{i'} \cdot (1-\alpha)}{L_P^{ii'} \cdot L_S^{ii'}} + \sum_{j \in \Omega} \frac{P_j}{L_P^{ij} \cdot L_S^{ij}} + N_0 \cdot W} \tag{4.6}$$

donde i' y Ω' hacen referencia a los usuarios en la celda de referencia i distintos al usuario de interés. En este caso, en $P_{i'}$ y P_j también se incluye la potencia de estación base destinada a otros canales distintos del canal HS-DSCH (*High Speed Downlink Shared CHannel*) de HSDPA (ver apéndice A). En cada estación base se ha considerado que un 80 % de la potencia total de transmisión para HSDPA está dedicada al canal HS-DSCH. El 20 % restante corresponde a otros canales de control y señalización que también contribuyen a la componente de interferencia intracelda.

En la estimación de interferencias se han considerado las estaciones base pertenecientes a la primera corona de celdas interferentes. Para GPRS y EDGE dicha corona está formada por las seis celdas cocanales más cercanas a la celda interferida, mientras que para HSDPA dicha corona está formada por las seis celdas vecinas a la celda interferida. Tomando como referencia la celda número 10 de la figura 4.2, la primera corona en GPRS y EDGE estaría formada por las celdas 1, 2, 9, 11, 18 y 19 (factor de reutilización igual a 3), mientras que en HSDPA estaría formada por las celdas 5, 7, 8, 13, 14 y 16 (factor de reutilización igual a 1).

4.2.4.3. Modelo de tráfico

Este elemento es el responsable de generar el tráfico de datos asociado a cada terminal móvil del sistema. En principio, dicho elemento debería haber sido implementado dentro de la entidad *Estación base*, ya que en la plataforma de simulación la transmisión de datos se simula en sentido descendente y, por lo tanto, el tráfico de datos se genera en la estación base. Sin embargo, se ha elegido esta ubicación por motivos de optimización del código.

En la literatura se han propuesto diversos métodos para modelar el tráfico de una red, los cuales pueden clasificarse en dos grupos, cada uno de los cuales se basa en una aproximación diferente. Una primera aproximación consiste en monitorizar y capturar una sesión de tráfico en una red real. Posteriormente, durante la simulación, se reproduce la sesión previamente monitorizada. Una aproximación alternativa consiste en intentar obtener modelos matemáticos que describan con aceptable verosimilitud estadística cada uno de los parámetros representativos del tipo de tráfico considerado. Posteriormente, estos modelos se utilizan para generar el tráfico durante la simulación. Los modelos de tráfico implementados en la plataforma de simulación corresponden a este segundo grupo. Concretamente, en la plataforma de simulación se han implementado modelos de tráfico para navegación web, correo electrónico y vídeo H.263 en tiempo real a diferentes velocidades binarias medias en múltiplos enteros de 16 kbps. A continuación se describe cada uno de estos tres modelos.

Modelo de tráfico de navegación web La navegación web hace uso del protocolo HTTP (*HyperText Transfer Protocol*) en la capa de aplicación, bajo la cual reside TCP (*Transmission Control Protocol*) como protocolo de la capa de transporte. Cuando un usuario solicita la descarga de una página web, está solicitando la transferencia de los diversos elementos u objetos que la integran: texto, imágenes, botones, listas desplegables, etc. Para la transferencia de estos objetos, el protocolo HTTP establece conexiones TCP. En las versiones HTTP/0.9 y HTTP/1.0 [167] se establece una conexión TCP diferente para cada objeto de la página web transmitida, mientras que en la versión HTTP/1.1 [168] se permite que todos los objetos de la página sean transmitidos utilizando una única conexión TCP. En el primer caso, se habla de *conexiones TCP por objeto*, mientras que en el segundo caso, de *conexiones TCP por página*.

El modelo de tráfico web implementado [169] está basado en el uso de conexiones TCP por objeto y su funcionamiento se ilustra en la figura 4.4. Durante una sesión web un usuario suele visitar varias páginas web diferentes. Al realizar la solicitud de una página web se inicia la fase de transferencia, que concluye cuando el usuario ha recibido por completo la página solicitada. Entonces el usuario emplea cierto tiempo en leer la información contenida en la página (fase de inactividad) hasta que hace clic sobre un vínculo de la página o escribe una nueva dirección, solicitando una nueva página web e iniciando así una nueva fase de transferencia. Puesto que el modelo web utilizado se basa en el uso de conexiones TCP por objeto, durante la fase de transferencia existen periodos activos y periodos inactivos. Los periodos activos corresponden a los intervalos de tiempo durante los cuales se está transmitiendo uno de los objetos que componen la página web, mientras que los periodos inactivos corresponden a los intervalos de tiempo en los que se cierra una conexión TCP y se establece una nueva para transmitir un nuevo objeto de la página web. La tabla 4.2 muestra cómo se modelan los parámetros representativos del tráfico web. El periodo activo en la fase de transferencia no aparece de forma explícita pero éste depende del tamaño del objeto.

Modelo de tráfico de correo electrónico El modelo de tráfico de correo electrónico implementado [170] supone que los correos se encuentran almacenados en un servidor de correo electrónico. Por lo tanto, cuando un usuario accede a su cuenta de correo a través del terminal móvil, se descarga solamente la cabecera de los correos almacenados en el servidor. Si el usuario decide leer un correo concreto entonces es cuando se produce la descarga desde el servidor. Este proceso de transferencia corresponde al periodo activo, después del cual el usuario emplea cierto tiempo en leer el correo. Desde que se termina de transferir un correo electrónico hasta que se solicita y comienza la transmisión de uno nuevo transcurre un intervalo de tiempo que recibe el nombre de periodo inactivo, después del cual comienza un nuevo periodo de actividad en el que se produce la descarga del segundo correo electrónico y así sucesivamente. La tabla 4.3 muestra los parámetros representativos del modelo. Nuevamente, el periodo activo no aparece de forma explícita en la tabla pero éste depende del tamaño del correo electrónico. Observar que el tamaño del correo electrónico se modela de manera *bimodal*, ya que el tamaño de un correo electrónico variará de forma significativa en función de si contiene archivos adjuntos o no.

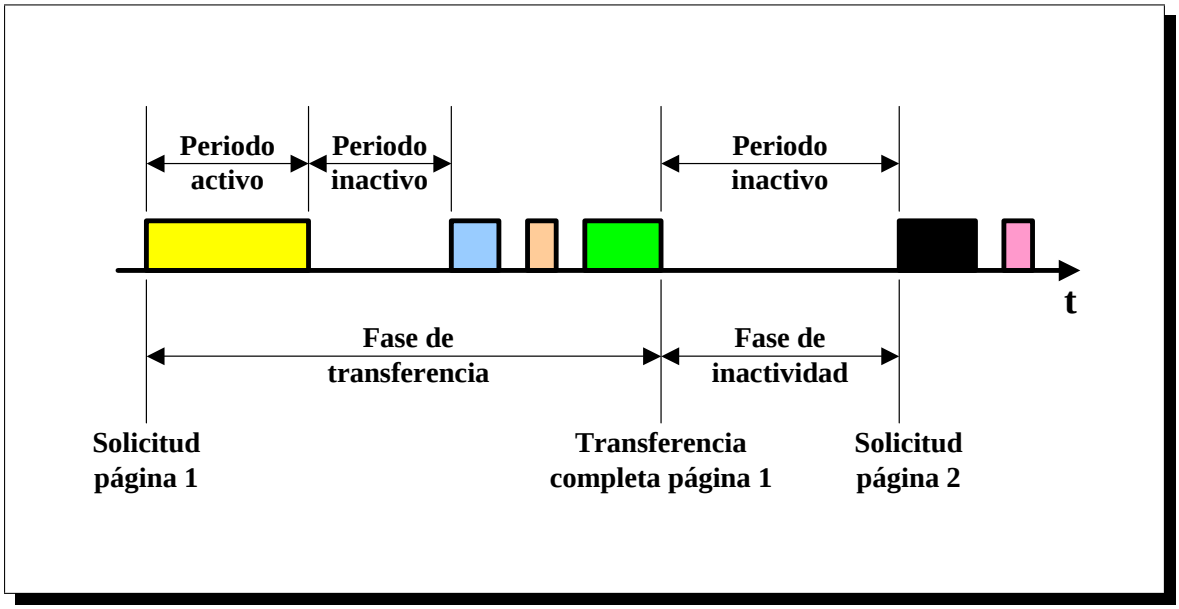


Figura 4.4: Modelo de tráfico de navegación web.

Parámetro representativo	Distribución matemática	Función de distribución de probabilidad (PDF)	Valor de las constantes
Tamaño de los objetos	Pareto	$f(x) = \frac{\alpha k^\alpha}{x^{\alpha+1}}$	$k = 1000$ $\alpha = 1{,}0$
Periodo inactivo, fase de transferencia	Weibull	$f(x) = \frac{b \cdot x^{b-1}}{a^b} e^{-(x/a)^b}$	$a = 1{,}46$ $b = 0{,}382$
Periodo inactivo, fase de inactividad	Pareto	$f(x) = \frac{\alpha k^\alpha}{x^{\alpha+1}}$	$k = 1$ $\alpha = 1{,}5$
Número de elementos por página web	Pareto	$f(x) = \frac{\alpha k^\alpha}{x^{\alpha+1}}$	$k = 1$ $\alpha = 2{,}43$

Tabla 4.2: Parámetros representativos del modelo de tráfico de navegación web.

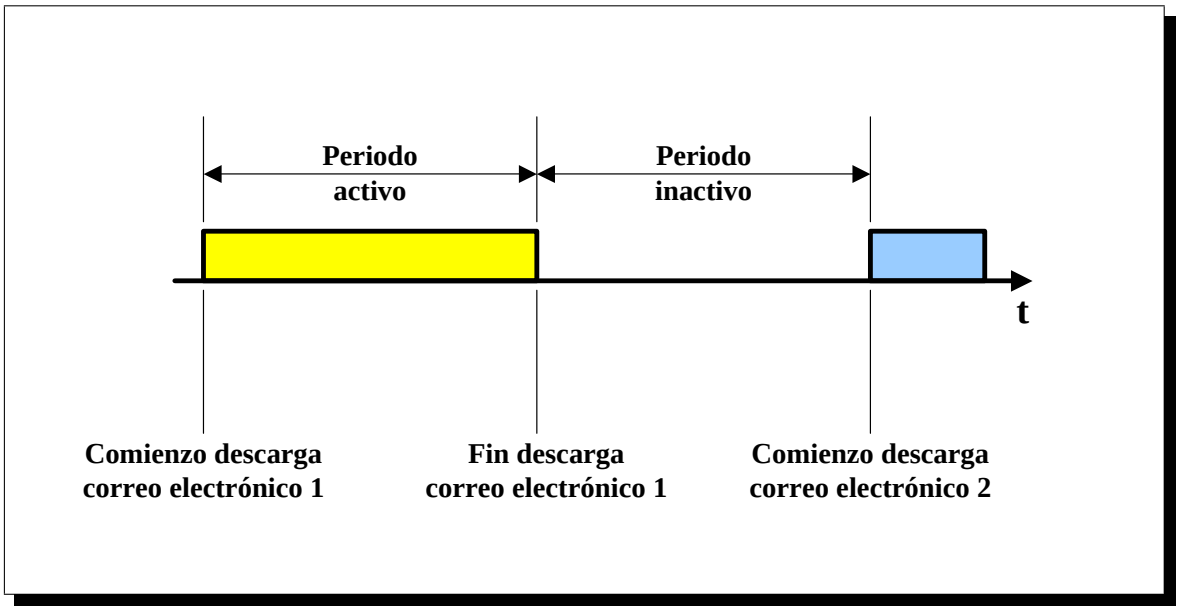


Figura 4.5: Modelo de tráfico de correo electrónico.

Parámetro representativo	Distribución matemática	Función de distribución acumulativa (CDF)	Valor de las constantes
Tamaño del correo	Weibull	$F(x) = \begin{cases} 1 - \exp\left[-e^{-k_1 x^{c_1}}\right], F(x) \leq 0{,}5 \\ 1 - \exp\left[-e^{-k_2 x^{c_2}}\right], F(x) > 0{,}5 \end{cases}$	$k_1 = 17{,}6$ $k_2 = 3{,}61$ $c_1 = 2{,}04$ $c_2 = 0{,}37$
Periodo inactivo	Pareto	$F(x) = 1 - \left(\frac{k}{x}\right)^{\alpha}$	$k = 30$ $\alpha = 1{,}5$

Tabla 4.3: Parámetros representativos del modelo de tráfico de correo electrónico.

Modelo de tráfico de vídeo H.263 Resulta fácil demostrar que para la transmisión de información de vídeo en una red de comunicaciones móviles es imprescindible la aplicación de algún método de codificación que permita conseguir elevados niveles de compresión ya que, de lo contrario, se requerirían tasas de transmisión varios órdenes de magnitud por encima de las que cualquier sistema de comunicaciones móviles puede proporcionar [171]. En la actualidad existen varios estándares de compresión de vídeo aunque sus características y modo de funcionamiento son bastante similares [172]. Los métodos de codificación estandarizados están dominados por dos familias [173]: la familia H.26x de la ITU (*International Telecommunication Union*) y la familia MPEG de la ISO (*International Standardization Organization*). El éstandar de codificación H.263 [174], que surgió con el objetivo de mejorar las prestaciones ofrecidas por su predecesor [175] (el estándar H.261), permite relaciones de compresión del orden de 100:1 o incluso 200:1 con una calidad de imagen aceptable [176], reduciendo la velocidad binaria requerida a valores en el rango de 16 a 300 kbps aproximadamente [177]. La posibilidad de transmitir con estas velocidades tan reducidas hace que la codificación H.263 sea sumamente atractiva para la transmisión de vídeo en un sistema de comunicaciones móviles [178], motivo por el cual se ha considerado el estándar H.263 para la plataforma de simulación.

Un codificador de vídeo H.263 tiene dos modos básicos de funcionamiento [174]: modo *Intra* y modo *Inter*. La secuencia de bits obtenida a la salida del codificador en el primer modo se denomina trama I, mientras que en el segundo modo recibe el nombre de trama P. En una trama de tipo I se halla codificada la información de una imagen completa, de forma que a partir de una trama I es posible reconstruir una imagen entera de la secuencia de vídeo. Se dice en este caso que en modo *Intra* la imagen se codifica sin referencia a ninguna otra imagen. En una trama de tipo P, en cambio, la imagen procesada se codifica haciendo referencia a la imagen predecesora. Es decir, cuando el codificador funciona en modo *Inter* sólo se codifican aquellas regiones de la imagen que varían respecto a la imagen anterior. Las zonas de la imagen que no varían no son codificadas, ya que el receptor recibió esa información en la trama anterior.

En una secuencia de vídeo H.263, la primera trama de vídeo transmitida siempre ha de ser necesariamente una trama I. Tras esta primera trama, el resto de tramas de vídeo transmitidas pueden ser de tipo P. Con la transmisión de tramas P se consigue reducir el ancho de banda necesario para transmitir la secuencia de vídeo, ya que las tramas P, al codificar sólo una parte de la imagen transmitida, poseen un tamaño mucho más reducido que las tramas I, las cuales contienen una imagen completa. Sin embargo, la imagen de una trama P puede ser decodificada de forma incorrecta, incluso habiéndose recibido la trama P correspondiente sin error, en caso de que la trama anterior hubiera sufrido un error de transmisión, ya que una trama P se codifica y decodifica haciendo referencia a la imagen predecesora. Así pues, el error de transmisión no sólo afecta a la imagen dañada sino también a la imagen siguiente, la cual servirá, a su vez, de referencia para la que se decodifique a continuación, con lo que el error se propagará nuevamente. Un error de transmisión en una imagen de vídeo puede propagarse durante tiempo suficiente como para que el usuario perciba un efecto desagradable [179]. La acumulación de errores a lo largo de la secuencia de

vídeo da como resultado una calidad final bastante pobre para el usuario, aun cuando los errores individuales experimentados sean poco significativos. Por ello, durante la transmisión se insertan tramas I cada cierto tiempo. De esta forma, si ocurre un error en alguna de las tramas transmitidas y éste afecta a varias tramas P posteriores, la distorsión que el error produce en la imagen dejará de propagarse en el momento en que se reciba una trama I, ya que ésta se decodifica de forma independiente. La frecuencia de transmisión de tramas I se denomina frecuencia de refresco y su valor establece un compromiso entre calidad de imagen y ancho de banda requerido por la secuencia de vídeo. Si se aumenta la frecuencia de refresco, entonces mejora la calidad de imagen que experimenta el usuario pero a cambio de incrementar el ancho de banda requerido.

El estándar H.263 contempla la posibilidad de aplicar cuatro modos opcionales a la codificación de vídeo H.263 básica (ver [176]). De estos cuatro modos, el modelo empleado implementa el modo de tramas PB. El modo de tramas PB permite incrementar la tasa de imágenes por unidad de tiempo sin un aumento significativo de la tasa binaria requerida [175], ya que en la misma trama se transmite la información relativa a dos imágenes de vídeo codificadas. Una trama PB consiste en una imagen P (codificada en modo *Inter*) y una imagen B. Las imágenes B, al igual que las imágenes P, se codifican haciendo referencia a otras imágenes. La diferencia está en que la imagen B se codifica haciendo referencia, no sólo a la imagen anterior como es el caso de las imágenes P, sino también a la imagen posterior. Para generar una trama PB es necesario que el transmisor espere hasta poseer las dos imágenes que se van a codificar en dicha trama. En el receptor, primero se reconstruye la imagen P a partir de la imagen de la trama de vídeo anterior y luego se reconstruye la imagen B a partir de la imagen P recién reconstruida y de la imagen de la trama anterior. Cuando concluye la reconstrucción de las dos imágenes, primero se muestra la imagen B y luego la imagen P [175]. La figura 4.6a muestra de forma gráfica el proceso de codificación H.263.

El modelo de tráfico de vídeo H.263 empleado en el simulador [180, 181] contempla los tres tipos de tramas de vídeo del estándar H.263: I, P y PB. El modelo captura de forma precisa las características de cada uno de estos tres tipos de trama. Las propiedades y variables manejadas por el modelo son el tamaño de las tramas de vídeo, su duración (ver figura 4.6b), la correlación existente entre el tamaño de una trama y su duración y la frecuencia o probabilidad de transición entre los diferentes tipos de tramas. El modelado se implementa en dos niveles diferentes.

El primer nivel de modelado determina el tipo de trama a transmitir. Las tramas I se generan a intervalos de tiempo regulares mientras que la generación de tramas P y PB se determina a partir de una cadena de Markov como la mostrada en la figura 4.7. La cadena de Markov se utiliza de la siguiente forma. Estando en el estado x se genera un número aleatorio X_0 a partir de una distribución de probabilidad uniforme entre 0 y 1. Si $X_0 \leq P_{x-x}$ entonces se permanece en el estado x, mientras que si $X_0 > P_{x-x}$ entonces se produce una transición hacia el estado y.

Una vez establecido el tipo de trama x a generar, el segundo nivel de modelado determina su tamaño T_x y su duración D_x mediante las siguientes transformaciones:

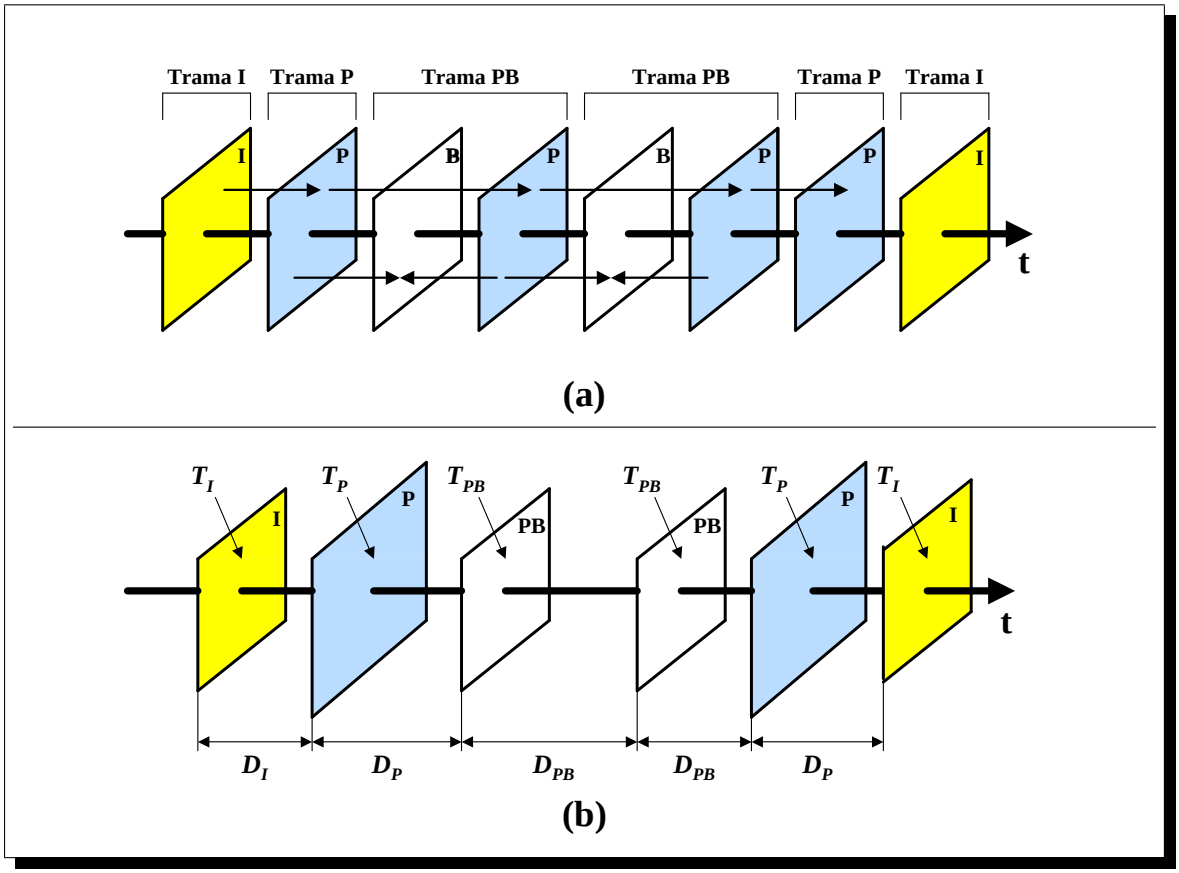


Figura 4.6: Codificación de vídeo H.263. (a) Ilustra el proceso de codificación H.263. I, P y B representan el tipo de imagen. (b) Ilustra las características del tráfico de vídeo H.263. I, P y PB representan el tipo de trama, cada una de las cuales se caracteriza por un tamaño (T_I, T_P, T_{PB}) y una duración (D_I, D_P, D_{PB}).

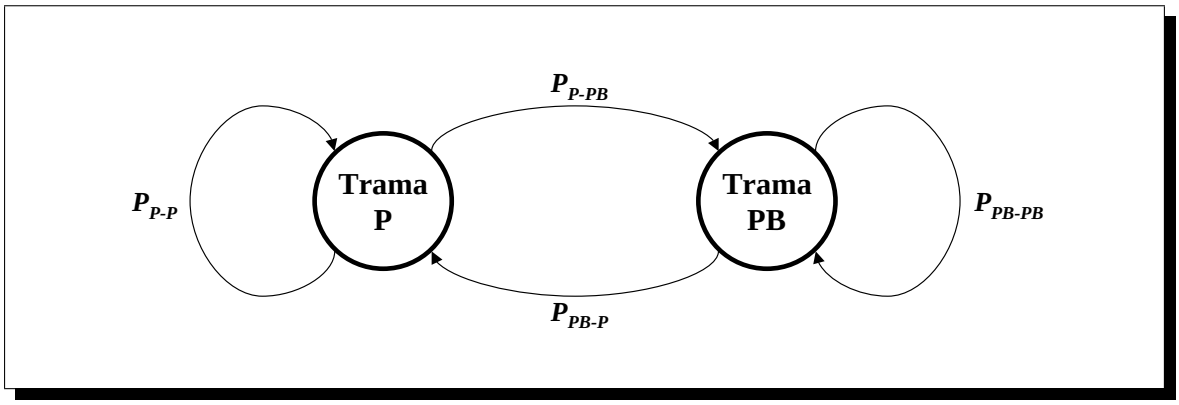


Figura 4.7: Modelo de Markov para la transición entre tramas H.263. Una cadena de Markov representa un modelo compuesto por un conjunto finito de estados discretos. En su representación gráfica, los estados se encuentran unidos entre sí mediante flechas. Cada flecha representa la transición entre dos estados y sobre ella se indica la probabilidad de que dicha transición ocurra. P_{x-y} denota la probabilidad de que se produzca una transición desde el estado x hacia el estado y.

$$T_x = F_T^{-1}\left(F_X(X_0)\right) \quad (4.7)$$
$$D_x = F_D^{-1}\left(F_T(T_x)\right) \quad (4.8)$$

donde X_0 es un número aleatorio generado a partir de una distribución normal con una media y varianza determinada que dependen del tipo de trama x generada, $F_X(\cdot)$ es la función de distribución acumulativa o CDF (*Cumulative Distribution Function*) de X_0 y $F_T^{-1}(\cdot)$ y $F_D^{-1}(\cdot)$ son, respectivamente, las inversas de las CDFs obtenidas a partir del tamaño y la duración de las tramas de una secuencia de vídeo real. En la expresión 4.8 se observa que el tamaño de la trama generada interviene a la hora de determinar su duración. En efecto, existe cierta correlación entre el tamaño de una trama de vídeo y su duración [180, 181], la cual queda modelada mediante este procedimiento.

Las CDFs anteriores, las probabilidades de transición de la figura 4.7 y el valor de otros parámetros utilizados por el modelo empleado se obtienen a partir de una secuencia de vídeo real codificada con H.263. Dicha secuencia de vídeo, con una duración aproximada de una hora, ha sido obtenida de [182], donde se encuentran disponibles diversas secuencias de vídeo reales (con una duración aproximada de una hora) codificadas con H.263 para tasas binarias medias de 16 kbps, 64 kbps y 256 kbps. Las secuencias para otras velocidades binarias distintas de las anteriores se obtienen a partir de una de ellas multiplicando el tamaño de las tramas y manteniendo constante su duración. Así, pueden obtenerse secuencias de vídeo a 32 kbps, 128 kbps y 512 kbps a partir de las secuencias de 16 kbps, 64 kbps y 256 kbps, respectivamente, multiplicando por dos el tamaño de las tramas de vídeo sin modificar su duración.

Por último, cabe señalar que la transmisión de vídeo H.263 en la plataforma de simulación se realiza en tiempo real. Esto significa que cada vez que se genera una trama de vídeo H.263 dicha trama lleva asociado un plazo de tiempo (su duración) dentro del cual debe completarse la transmisión de la misma. Una vez finalizado dicho plazo se genera la siguiente trama de vídeo H.263 y así sucesivamente. Si en el momento de generarse una trama de vídeo H.263 no ha finalizado la transmisión de la trama previa, entonces la información de la trama previa pendiente de ser transmitida es descartada para que comience de inmediato la transmisión de la nueva trama de vídeo.

4.2.4.4. Módulo CRRM

Este elemento es el responsable de las funciones CRRM implementadas en la plataforma de simulación. La función CRRM de interés en el presente trabajo es la de decidir en cada momento la RAT óptima por la cual deben transmitirse los paquetes de datos del usuario. Por lo tanto, es en el módulo CRRM donde se implementan los algoritmos evaluados en el presente trabajo. La implementación de este módulo se ha realizado de manera que ofrezca flexibilidad. Un usuario puede cambiar de tecnología de acceso en cada nuevo bloque de datos transmitido (objeto web, correo electrónico o trama de vídeo H.263), cada cierto intervalo de tiempo o al inicio de una sesión.

4.2.5. Control de enlace

En este elemento, común a las entidades *Estación base* y *Terminal móvil*, se implementan todos los aspectos específicos de cada RAT, reproduciendo de manera fiel el proceso de transmisión de datos a través de la interfaz radio de cada una de ellas. En el apéndice A se resumen las principales características de cada RAT.

En primer lugar, en el elemento *Control de enlace* se implementan los aspectos relativos al acceso múltiple FDMA/TDMA de GPRS y EDGE según las especificaciones ETSI GSM 05.01 y ETSI GSM 05.02, así como los relativos al acceso múltiple CDMA de HSDPA en su componente FDD según se indica en 3GPP TR 25.858.

También se han incluido las características de los diferentes modos de transmisión disponibles en cada RAT y se ha modelado su funcionamiento adaptativo mediante la técnica conocida como *Link Adaptation* (LA) o *Adaptive Modulation and Coding* (AMC), consistente en seleccionar de forma periódica el modo de transmisión que resulta óptimo para la calidad experimentada en el enlace radio, siguiendo cierto criterio predefinido. En el caso de tráfico web y correo electrónico se ha aplicado el criterio ampliamente extendido consistente en seleccionar aquel modo de transmisión que, para la calidad experimentada, ofrece la mayor cantidad de bits correctamente transmitidos por unidad de tiempo (*throughput*) [183]. En el caso de tráfico de vídeo H.263 se ha aplicado el algoritmo descrito en [184–186], que mejora al anterior en varios aspectos de vital importancia para servicios en tiempo real.

Para evaluar correctamente el rendimiento de las técnicas CRRM bajo estudio así como para estimar adecuadamente la calidad de servicio final que perciben los usuarios del mismo, la plataforma de simulación modela la transmisión de datos con un elevado nivel de resolución temporal e incluye los efectos de la transmisión radio a través de diversas LUTs, las cuales se emplean para decidir si los bloques de datos transmitidos se reciben correctamente o se reciben con error. El procedimiento empleado para cada RAT se detalla a continuación.

En el caso de GPRS y de EDGE, la información se transmite a través de la interfaz radio en forma de bloques radio (*radio blocks*) con una duración de 20 ms. Cada bloque radio se divide en cuatro ráfagas normales (*normal bursts*) cada una de las cuales se acomoda en un *time-slot* a lo largo de cuatro tramas TDMA consecutivas (ver apéndice A). Cada vez que se transmite una ráfaga normal se obtiene el valor de CIR experimentado en su transmisión. Cuando se completa la transmisión de un bloque radio, se promedian los valores de CIR correspondientes a las cuatro ráfagas normales requeridas para su transmisión. Dicho promediado se realiza calculando la media geométrica de los cuatro valores de CIR, es decir, la media aritmética de los valores de CIR expresados en decibelios. Con ello se obtiene un valor de CIR representativo de la calidad de transmisión experimentada por el bloque radio. Para decidir si dicho bloque se recibe correctamente o se recibe con error se recurre a las LUTs. A partir del valor de CIR medio obtenido anteriormente y de la curva correspondiente al modo de transmisión empleado, se obtiene un valor de BLER tal y como se muestra en la figura 4.8. El

valor de BLER obtenido, denotado como X_{CIR} en la figura, se utiliza para decidir si el bloque se recibe o no con error. El procedimiento empleado para tomar dicha decisión [187] consiste en generar un número aleatorio X_0 a partir de una variable aleatoria uniformemente distribuida entre 0 y 1. Si $X_0 \leq X_{CIR}$, se considera que el bloque radio ha sido recibido con error. Si $X_0 > X_{CIR}$, entonces se considera que el bloque radio ha sido correctamente recibido. El procedimiento completo se ilustra en la figura 4.9.

En el caso de HSDPA, la información se transmite por la interfaz radio en forma de bloques de transporte (*transport blocks*) con una duración de 2 ms. Dicho intervalo de tiempo se divide en tres *slots* de 667 μs. Para cada bloque de transporte se obtiene el valor de SIR en cada uno de sus tres *slots* y, realizando una media geométrica, se obtiene un valor de SIR medio representativo de la calidad experimentada por dicho bloque. En HSDPA, cuando un bloque se recibe con error, éste no es descartado sino que se almacena en el *buffer* receptor y posteriormente se combina con las sucesivas retransmisiones del mismo bloque utilizando los métodos *Chase Combining* (CC) o *Incremental Redundancy* (IR). Por este motivo el procedimiento descrito anteriormente no resulta adecuado para HSDPA y se hace necesario emplear otro procedimiento.

En primer lugar se combina el valor de SIR del último bloque recibido con los valores de SIR de las transmisiones previas del mismo bloque (en caso de que las haya), obteniendo así un valor de SIR efectivo representativo de la calidad experimentada por la información disponible en el *buffer* de recepción a lo largo de las diversas transmisiones. Esta combinación de los valores de SIR pretende modelar los efectos de los métodos CC o IR y se ha implementado según se describe en [188]. Después de $K+1$ transmisiones de un mismo bloque, incluyendo la transmisión actual y K transmisiones previas, el valor de SIR efectivo viene dado por [188]:

$$SIR_{EFECTIVO} = \varepsilon^K \cdot \eta(M,R) \cdot \sum_{n=0}^{K} SIR_{MEDIO}(t-n) \tag{4.9}$$

donde los valores de SIR están expresados en magnitud lineal (no en decibelios). Este modelo general es aplicable tanto a CC como a IR. El parámetro ε representa la eficiencia del proceso de combinación. Este valor es siempre inferior a la unidad, indicando que existen ciertas pérdidas por combinación, de entre unos 0.2 y 0.3 dB aproximadamente [188]. El parámetro $\eta(M,R)$ representa la ganancia de IR sobre CC. Cuando se aplica CC su valor es siempre igual a la unidad. Cuando se aplica IR su valor es igual a la unidad en la transmisión inicial ($K=0$). En las sucesivas retransmisiones con IR ($K \geq 1$), el valor de $\eta(M,R)$ es superior a la unidad y depende del orden de modulación M ($M=4$ para QPSK y $M=16$ para 16-QAM) y de la tasa efectiva R del código de canal empleado. Observar que en la primera transmisión de un bloque ($K=0$) el valor de SIR efectivo se corresponde, tal y como cabe esperar, con el valor de SIR experimentado en dicha transmisión. Los parámetros del modelo se muestran en las tablas 4.4 y 4.5 (las cuales han sido extraídas de [188]), para entornos de propagación pedestre (Ped-A) y vehicular (Veh-A). La tasa de código R de algunos modos de transmisión de HSDPA no aparece en dichas tablas. Para aquellos valores de R no indicados explícitamente, el valor de $\eta(M,R)$ utilizado se ha obtenido por interpolación lineal a partir de los valores inmediatamente superior e inferior.

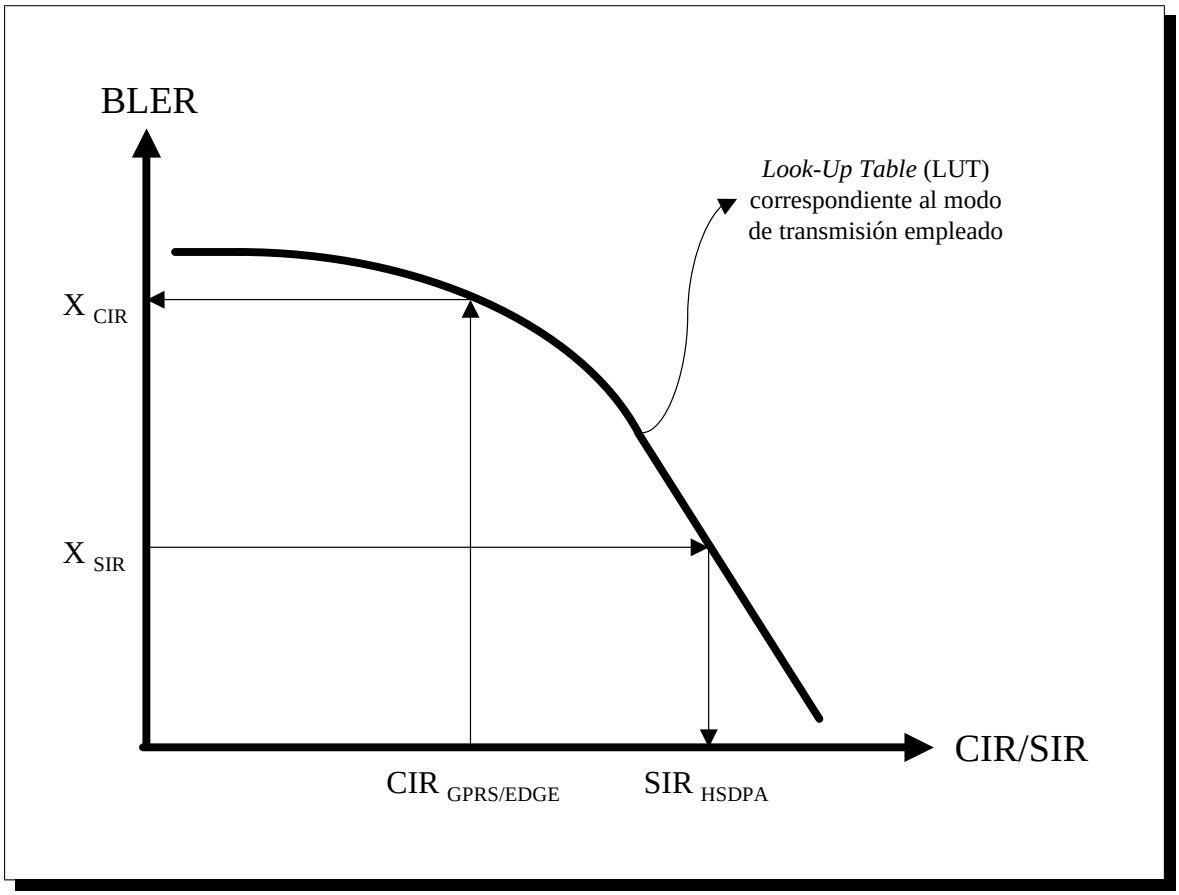


Figura 4.8: Modo de utilización de las *Look-Up Tables* (LUTs).

Parámetro	R	Ped-A	Veh-A
$\eta(4, R)$	0.20	1.03	1.03
	0.30	1.01	1.01
	0.40	1.03	1.04
	0.50	1.07	1.11
	0.60	1.14	1.20
	0.70	1.24	1.35
	0.80	1.42	1.59
ε	-	0.93	0.95

Tabla 4.4: Parámetros QPSK (3 km/h).

Parámetro	R	Ped-A	Veh-A
$\eta(16, R)$	0.50	1.35	1.46
	0.55	1.42	1.54
	0.60	1.48	1.59
	0.65	1.62	1.71
	0.70	1.75	1.81
	0.75	1.87	1.93
ε	-	0.92	0.93

Tabla 4.5: Parámetros 16-QAM (3 km/h).

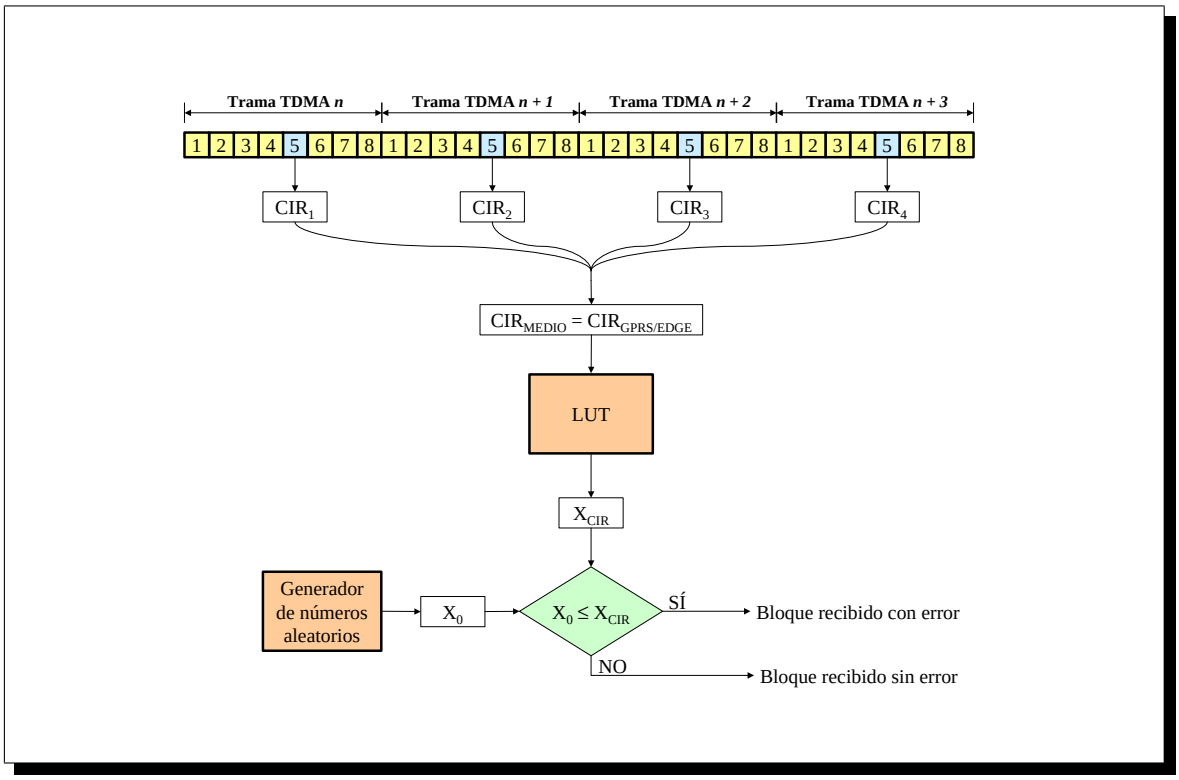


Figura 4.9: Modelado de la calidad del canal en GPRS y EDGE.

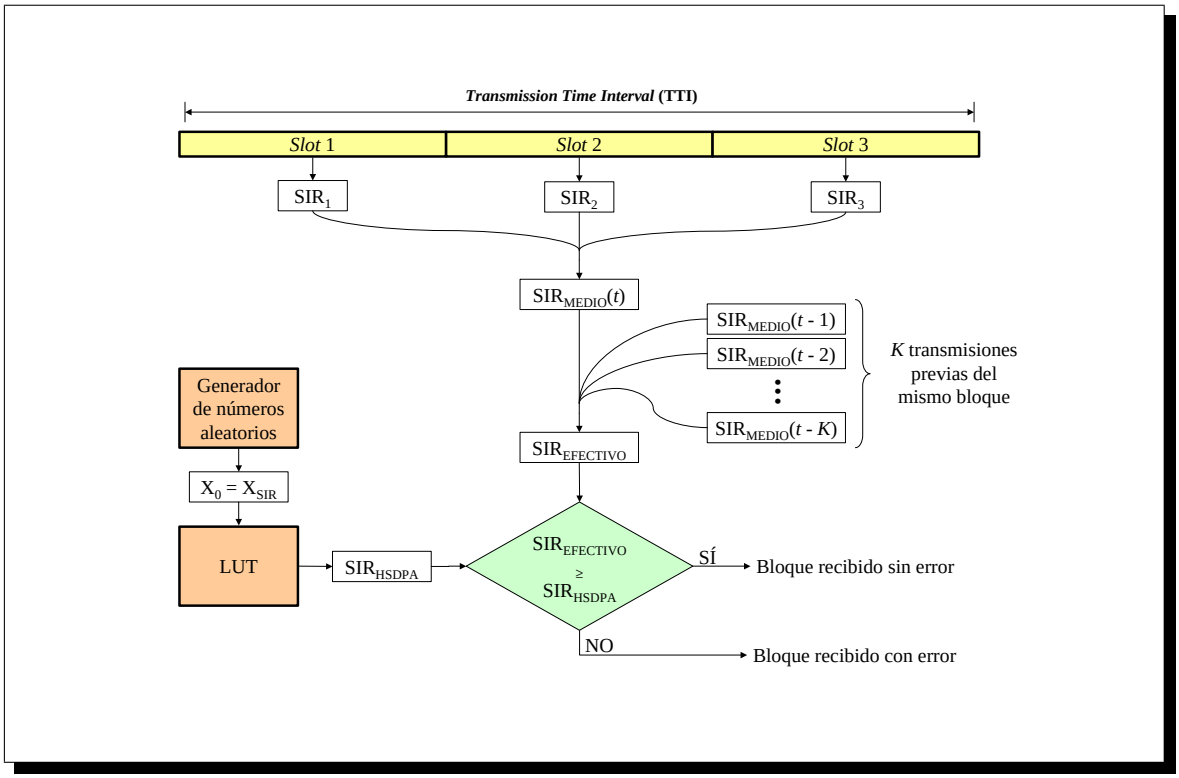


Figura 4.10: Modelado de la calidad del canal en HSDPA.

Una vez que se dispone del valor de SIR efectivo tras la última transmisión, se decide si el bloque ha sido recibido o no con error siguiendo el procedimiento descrito en [189]. Dicho procedimiento consiste en generar, antes de la primera transmisión de un nuevo bloque de transporte, un número aleatorio X_0 a partir de una distribución de probabilidad uniforme entre 0 y 1. A partir del valor X_0 obtenido y de la curva correspondiente al modo de transmisión que se va a emplear para transmitir dicho bloque, se obtiene un valor de SIR tal y como se muestra en la figura 4.8. En dicha figura, el valor de X_0 sería el indicado como X_{SIR} y el valor de SIR obtenido sería el indicado como SIR_{HSDPA}. El valor SIR_{HSDPA} así obtenido, se establece como valor mínimo de SIR efectivo que debe obtenerse en recepción para considerar que el bloque ha sido recibido correctamente. Si el valor de SIR experimentado en la primera transmisión del bloque es inferior a SIR_{HSDPA}, se considera que el bloque se recibe con error y se procede a su retransmisión. En la siguiente retransmisión, el valor de SIR experimentado se combina con el de la transmisión inicial para obtener un nuevo valor de SIR efectivo. Si dicho valor continúa siendo inferior a SIR_{HSDPA} se repite el proceso y así sucesivamente hasta que el valor de SIR efectivo resultante para un determinado bloque es mayor o igual que SIR_{HSDPA} o hasta que se cumple el número máximo de retransmisiones permitidas. Es importante señalar que el valor de SIR_{HSDPA} se determina antes de la primera transmisión de un bloque dado y dicho valor es utilizado en la comparación con el valor de SIR efectivo mientras dura la transmisión de dicho bloque.

Así pues, para cada bloque de transporte se emplea el mismo valor de SIR_{HSDPA} en todas las transmisiones de dicho bloque. Una opción alternativa consistiría en determinar un valor aleatorio de SIR_{HSDPA} en cada transmisión, de manera similar a como se realiza para GPRS y EDGE. No obstante, tal y como se señala en [189], esta forma de proceder no modelaría de forma correcta el efecto de las retransmisiones en HSDPA. Considérese, por ejemplo, que en la transmisión inicial de un bloque el valor de SIR experimentado es inferior a SIR_{HSDPA} y, por lo tanto, se considera recibido con error. Supongamos que en la siguiente transmisión de dicho bloque el valor de SIR experimentado es extremadamente pobre, lo cual implica que dicha retransmisión no aporta una mejora significativa al valor de SIR obtenido en la transmisión inicial. Es decir, el valor de SIR efectivo después de la retransmisión permanece aproximadamente igual que después de la transmisión inicial. Esto significa que si no pudo decodificarse correctamente la transmisión inicial, tampoco debería poder hacerse con la información disponible después de la retransmisión efectuada. No obstante, si después de la retransmisión se vuelve a generar un nuevo valor aleatorio de SIR_{HSDPA}, cabe la posibilidad de que la información se considere recibida correctamente después de la retransmisión, lo cual sería incorrecto. Por ello, el valor de SIR_{HSDPA} utilizado en el proceso de decisión permanece constante durante todas las transmisiones asociadas a un mismo bloque. Esta manera de proceder, por lo tanto, se aproxima más al comportamiento de un sistema real. El procedimiento completo se ilustra en la figura 4.10.

Por último, para la retransmisión de los bloques recibidos con error, se ha implementado al detalle los protocolos correspondientes. En GPRS y EDGE se emplea un protocolo ARQ de ventana deslizante (3GPP TS 04.60). En HSDPA se utiliza un protocolo ARQ de parada y espera con varios procesos simultáneos (3GPP TR 25.848).

4.3. Configuración

La tabla 4.6 muestra un resumen de la configuración empleada para efectuar las simulaciones llevadas a cabo en el presente trabajo. Tal y como se ha señalado en las secciones previas, el sistema simulado es un entorno macrocelular compuesto de 27 celdas omnidireccionales organizadas según una estructura de clúster compuesta por 3 celdas, en las cuales se estudia el enlace descendente, es decir, desde la estación base hacia el terminal móvil. Las celdas tienen un radio de 500 metros y la distancia mínima de separación al punto central, donde se haya la estación base, es de 10 metros.

El número de canales de transmisión disponibles en cada celda se ha establecido en 4 para cada una de las tecnologías de acceso radio consideradas en la plataforma de simulación. A efectos de estimación de interferencias, se supone que las frecuencias portadoras de los canales son tales que un canal activo de una determinada RAT no puede interferir sobre otro canal activo de una RAT distinta. Cuando un usuario solicita un canal de una RAT determinada, se le asigna uno escogido de forma aleatoria de entre los canales libres. En caso de no haber canales disponibles, el usuario pasa a una cola de espera. Para los usuarios en las colas de GPRS y EDGE se aplica una política de *scheduling* de tipo FCFS, mientras que para los usuarios en la cola de HSDPA se aplica una política de tipo *Round Robin* con un cuanto temporal de transmisión de 100 ms.

Los usuarios se desplazan por el interior de las celdas a una velocidad constante de 3 km/h según el modelo descrito en la sección 4.2.4.1. Cuando un usuario llega al borde de la celda se le fuerza a volver hacia atrás, de tal forma que el número de usuarios por celda permanece siempre constante. El *handover* horizontal, por lo tanto, no ha sido implementado. Tampoco se ha implementado el control de potencia. En el caso de HSDPA no resulta necesario ya que la transmisión del canal HS-DSCH siempre se realiza a potencia constante (ver apéndice A). En el caso de GPRS y EDGE, finalmente, ha sido desactivado. Las transmisiones, por lo tanto, se realizan a potencia constante. Para GPRS y EDGE la potencia de transmisión es de 30 dBm, mientras que para HSDPA se ha establecido en 37 dBm. En el caso de HSDPA, un 80 % de dicha potencia (36.03 dBm) se dedica al canal HS-DSCH, mientras que el 20 % restante se considera reservado para otros canales (de control y señalización) que contribuyen a generar interferencia tal y como se comentó en la sección 4.2.4.2. La potencia reservada para el canal HS-DSCH se divide a partes iguales entre el número de códigos disponibles para dicho canal. Por tanto, la potencia de transmisión por usuario es igual a 37 dBm + $10 \cdot \log_{10}(0{,}8) + 10 \cdot \log_{10}(1/4) = 30$ dBm, que es el valor utilizado para GPRS y EDGE.

Los efectos de la atenuación o *pathloss* se modelan según el conocido modelo Okumura-Hata en su extensión COST 231 (ver sección 4.2.2), utilizando una frecuencia portadora $f_c = 1800$ MHz para GPRS/EDGE y $f_c = 2000$ MHz para HSDPA. Para el resto de parámetros del modelo se han empleado los siguientes valores: $A = 46.3$, $B = 33.9$, $C = 44.9$, $h_b = 40$ m y $h_m = 1.5$ m. La expresión de $a(h_m)$ empleada es la correspondiente a pequeñas y medianas ciudades. El efecto sombra o *shadowing* se modela mediante una distribución log-normal con una desviación estándar de 6 dB y una distancia de decorrelación de 20 m (ver sección 4.2.2).

Parámetro	**GPRS**	**EDGE**	**HSDPA**
Entorno	Urbano macrocelular		
Enlace estudiado	Descendente (*downlink*)		
Celdas simuladas	27 (omnidireccionales)		
Factor de reutilización	3	3	1
Radio de las celdas	500 m (distancia mínima al centro = 10 m)		
Canales por celda	4	4	4
Canales por usuario	1	1	1
Asignación de canal	Aleatoria (se elige al azar uno disponible)		
Scheduling	FCFS	FCFS	*Round Robin*
Movilidad	Implementada (velocidad constante = 3 km/h)		
Handover horizontal	No implementado		
Control de potencia	No implementado		
Potencia de transmisión	30 dBm	30 dBm	37 dBm
Modelo de *pathloss*	Okumura-Hata COST 231		
Modelo de *shadowing*	Ley log-normal		
Densidad espectral de ruido térmico	-174 dBm/Hz		
Ruido térmico	-121 dBm	-121 dBm	-107 dBm
Protocolo ARQ	Ventana deslizante (rechazo selectivo)		Parada-y-espera
Características protocolo ARQ	Ventana de 64 bloques RLC, confirmación cada 16 bloques	Ventana de 192 bloques RLC, confirmación cada 32 bloques	4 procesos simultáneos, máximo 4 transmisiones
Periodo LA/AMC	60 ms	60 ms	2 ms
Periodo de decisión de RAT	10 segundos		

Tabla 4.6: Configuración de la plataforma de simulación.

Los valores de potencia de ruido térmico se han obtenido considerando una densidad espectral de ruido de -174 dBm/Hz y un ancho de banda de 200 kHz para los canales de GPRS/EDGE y de 5 MHz para los canales de HSDPA. Estos cálculos, sin ser rigurosos, resultan, no obstante, suficientes para los objetivos del presente trabajo.

Para los protocolos de retransmisión de GPRS y EDGE, basados en el principio de ventana deslizante con rechazo selectivo, el tamaño de ventana se ha establecido de acuerdo a la especificación 3GPP TS 04.60. El tamaño de ventana indicado para GPRS es de 64 bloques RLC, mientras que para EDGE el tamaño de ventana es variable y depende del número de canales que tenga asignados de manera simultánea el terminal móvil. Por ejemplo, en transmisión monocanal, es decir con un único canal, el tamaño máximo de ventana permitido en EDGE es de 192 bloques RLC. El periodo de confirmación, que determina la frecuencia con la cual el receptor envía confirmaciones al transmisor, se ha establecido en 16 bloques RLC para GPRS siguiendo las indicaciones de [190]. Para EDGE se ha escogido un periodo de confirmación de 32 bloques RLC. El protocolo de retransmisión de HSDPA, basado en el principio de parada-y-espera, se ha configurado para funcionar con 4 procesos HARQ simultáneos y realizar hasta un máximo de 4 transmisiones para un mismo bloque, incluyendo la transmisión inicial. Las sucesivas retransmisiones de un mismo bloque se combinan mediante *Chase Combining*.

Los protocolos de retransmisión anteriores han sido utilizados para tráfico de navegación web y de correo electrónico. Para tráfico de vídeo H.263 los protocolos anteriores han sido desactivados. En el caso de servicios *no sensibles al retardo*, como es el caso de la navegación web y el correo electrónico, el retardo de transmisión posee cierta relevancia aunque no es un factor crítico. En cambio, sí resulta de gran importancia la integridad de los datos transmitidos. Dada la importancia que tiene en este tipo de servicios la transmisión de información libre de errores y puesto que toleran cierto retardo, se han aplicado los protocolos de retransmisión para navegación web y correo electrónico. En el caso de servicios *sensibles al retardo*, como es el caso de la transmisión de vídeo H.263 en tiempo real, el uso de protocolos de retransmisión resulta bastante cuestionable, ya que la retransmisión de bloques erróneos supone un retardo de transmisión adicional, generalmente inaceptable para este tipo de servicios. Por este motivo los protocolos de retransmisión han sido desactivados para la transmisión de vídeo H.263. En GPRS y EDGE esto significa que los bloques recibidos con error son descartados y no se retransmiten. En el caso de HSDPA existe cierto matiz, ya que un bloque recibido con error no se descarta de inmediato sino que se retransmite hasta que se recibe correctamente o hasta que se cumple el número máximo de transmisiones. Una vez alcanzado el límite de transmisiones para un mismo bloque, entonces es finalmente descartado. Esta forma de proceder para servicios en tiempo real, también aplicada por ejemplo en [191], se basa en que es necesario descartar la información recibida con error si no se quiere incurrir en un retardo de transmisión excesivo, lo cual resulta sumamente perjudicial para este tipo de servicios. Sin embargo, dadas las elevadas tasas de transmisión de HSDPA, parece poco sensato descartar un bloque en su primera transmisión errónea. Por ello se permiten algunas retransmisiones antes de descartar un bloque definitivamente. Además, según señala [166], el BLER experimentado en HSDPA en entornos macrocelulares en la primera transmisión toma valores

habituales comprendidos entre el 30 % y el 60 %, lo cual aconseja nuevamente permitir algunas retransmisiones antes de que la información sea definitivamente descartada.

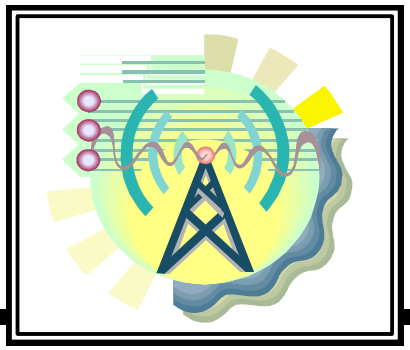
El periodo de actualización de LA/AMC, que determina la frecuencia con la cual se actualiza el modo de transmisión empleado, se ha establecido en 60 ms para GPRS y EDGE (valor típico), y en 2 ms para HSDPA (*fast link adaptation*).

Por último, el simulador ha sido configurado para que el algoritmo de decisión de selección de RAT se ejecute cada 10 segundos. En principio, el procedimiento lógico consistiría en ejecutar el algoritmo al inicio de la sesión de un servicio dado y mantener vigente la decisión tomada durante la duración de dicha sesión (típicamente superior a 10 segundos). No obstante, se ha considerado oportuno establecer este valor relativamente bajo para posibilitar que el algoritmo se ejecute en un elevado número de ocasiones y así obtener resultados en los cuales se aprecien mejor las consecuencias derivadas de las decisiones tomadas por el algoritmo bajo evaluación.

Antes de concluir esta sección merece la pena señalar que la configuración empleada para las simulaciones realizadas en este trabajo favorece la existencia de niveles de interferencia relativamente altos en el sistema. Por ejemplo, el valor seleccionado para el radio de las celdas, relativamente bajo, se ha establecido de manera acorde a los valores máximos habitualmente empleados en un sistema real para HSDPA; el bajo factor de reutilización cocanal empleado para GPRS y EDGE ha sido necesario para que el número total de celdas simuladas en el sistema resulte computacionalmente viable. Estas y otras decisiones dan lugar a la existencia de apreciables niveles de interferencia, lo cual repercute sobre el nivel de prestaciones ofrecido por cada una de las RATs del sistema. Esta circunstancia no afecta a la hora de evaluar comparativamente el rendimiento de los algoritmos estudiados ya que todos ellos se evalúan bajo las mismas condiciones. No obstante resultará importante tenerlo presente en el capítulo 5 a la hora de entender ciertos resultados cuando los algoritmos se analizan de manera individual.

4.4. Sumario

La evaluación del rendimiento de los algoritmos estudiados en este trabajo se lleva a cabo mediante simulación por ordenador. En el presente capítulo se ha presentado la plataforma de simulación desarrollada para este trabajo. Dicha plataforma, consistente en la integración de tres sofisticados y complejos simuladores a nivel de sistema de las tecnologías GPRS, EDGE y HSDPA, implementa fielmente las especificaciones de la interfaz radio de cada una de ellas y emula el proceso de transmisión radio con un elevado nivel de resolución temporal, permitiendo una evaluación precisa tanto del rendimiento de las técnicas CRRM estudiadas en este trabajo como de la calidad final que percibirán los usuarios en una red heterogénea de comunicaciones móviles.



CAPÍTULO 5

Rendimiento de los Algoritmos Propuestos

En este capítulo se presentan y analizan los resultados obtenidos en las simulaciones que se han realizado con la finalidad de evaluar el rendimiento de los algoritmos propuestos en el presente trabajo. Las simulaciones realizadas en este trabajo han sido íntegramente llevadas a cabo haciendo uso de la plataforma de simulación de redes heterogéneas descrita en el capítulo 4, la cual ha sido específicamente desarrollada como parte de este trabajo. Dicha plataforma, según se vio en el capítulo 4, está compuesta por tres avanzados y detallados simuladores a nivel de sistema de las tecnologías de acceso radio GPRS (*General Packet Radio Service*), EDGE (*Enhanced Data-rates for GSM/Global Evolution*) y HSDPA (*High Speed Downlink Packet Access*), esta última evolución del actual estándar de tercera generación UMTS (*Universal Mobile Telecommunications System*) calificada habitualmente como tecnología 3.5G. Esta plataforma de simulación ofrece un escenario heterogéneo compuesto por diferentes redes de acceso radio con distintas prestaciones, sobre el cual es posible emular diversos tipos de servicios con distintos requisitos de QoS. Esta plataforma, así pues, ofrece un escenario adecuado para evaluar el rendimiento de las técnicas CRRM que se proponen en el presente trabajo. Antes de entrar a analizar el rendimiento de los algoritmos propuestos resulta conveniente, no obstante, describir brevemente los parámetros que han sido considerados en este capítulo a la hora de evaluar el rendimiento de dichos algoritmos. Así pues, en primer lugar se discutirán los diversos parámetros de evaluación considerados en este capítulo y la manera en que éstos se definen, para a continuación analizar en profundidad durante las siguientes secciones el rendimiento de los algoritmos UBET y UBReQoS que se proponen en este trabajo. Por último se analizan las consecuencias derivadas de aplicar el método sugerido en este trabajo para estimar la calidad radio.

5.1. Parámetros y escenarios de evaluación

Puesto que cada RAT del sistema ofrece unas prestaciones notablemente diferentes, el rendimiento de los algoritmos y el nivel de QoS percibido por el usuario vendrán fuertemente determinados por la frecuencia con que se utilice cada una de las RATs del sistema. Así pues, el primer aspecto analizado en cada algoritmo es el porcentaje de selección o de utilización de cada RAT, tanto para cada tipo de servicio como a nivel global. El porcentaje de selección de una RAT determinada se define como el cociente entre el número de veces que se selecciona dicha RAT y el número de veces totales que se toma una decisión de selección de RAT (expresado en tanto por ciento).

Cuando una RAT determinada es seleccionada de forma bastante habitual, puede darse la situación, si el número de usuarios asignados a dicha RAT es lo suficientemente elevado, de que la cantidad de recursos disponibles no sea suficiente para atender las peticiones de todos los usuarios y varios de ellos tengan que permanecer en una cola a la espera de que se libere un canal. Para cuantificar el nivel de ocupación de las diversas colas del sistema, se introduce como parámetro de evaluación el número medio de usuarios en cola. Para calcular el valor de dicho parámetro, durante la simulación se ha muestreado con periodicidad suficiente el número de usuarios que se hayan esperando en cada una de las colas de espera de las RATs del sistema y, al final de la simulación, se han promediado los valores medidos en cada una de ellas.

La disponibilidad de diversas RATs en una localización determinada y la coordinación entre ellas según el concepto de red radio heterogénea ofrece diversas ventajas respecto al caso en el que las redes son operadas de manera individual y cada usuario está obligado a conectarse siempre a la misma RAT, tal y como se discutió en el capítulo 1. No obstante, una evaluación completa de las técnicas bajo estudio requiere considerar no sólo los aspectos positivos derivados de su aplicación sino también los costes asociados a su implantación. En el caso de interés que concierne al presente trabajo, el coste asociado a la implantación de las técnicas propuestas viene determinado por el nivel señalización introducida por los procesos de traspaso de usuarios entre las RATs del sistema (acentuado en este caso por la elevada frecuencia con que se toman las decisiones de selección de RAT, según muestra la tabla 4.6). La frecuencia de *handovers* o traspasos verticales debe ser especialmente considerada para lograr un proceso de selección de RAT estable, ya que es altamente indeseable que se produzcan constantes y frecuentes cambios de RAT de forma simultánea para un elevado número de usuarios. Por ello, también se ha considerado como parámetro de evaluación el número de cambios de RAT por minuto y por usuario. En principio podría haberse considerado simplemente el número de cambios de RAT por minuto. Sin embargo, el valor de dicho parámetro puede variar de manera significativa en función del número de usuarios presentes en el sistema. Dada esta circunstancia, y puesto que el rendimiento de los algoritmos se evalúa bajo diferentes condiciones de carga, se ha optado por considerar finalmente el número de cambios de RAT por minuto y por usuario. El valor de este parámetro se obtiene dividiendo el número de cambios de RAT ocurridos durante un intervalo de tiempo de un minuto entre el número total de usuarios sobre los cuales se ha contabilizado el número de cambios de RAT.

El nivel de QoS percibido por los usuarios del sistema se evalúa a través de varios parámetros. Un parámetro importante habitualmente utilizado para evaluar el nivel de QoS experimentado es el *throughput*, el cual se define como la cantidad de bits correctamente transmitidos por unidad de tiempo. Otro parámetro importante es el retardo de transmisión, que generalmente se define como el tiempo requerido para transmitir un determinado paquete de información. El valor absoluto del retardo de transmisión no permite en sí evaluar de forma equitativa las transmisiones de diferentes paquetes, ya que generalmente será mayor para aquellos paquetes de mayor tamaño. Por ello, en este trabajo se ha considerado en su lugar el retardo normalizado, el cual se define como el cociente entre el retardo de transmisión y el tamaño del paquete transmitido. Para calcular el valor de este parámetro se considera no sólo el tiempo de transmisión radio (contado a partir del instante en que se recibe el canal) sino también los posibles tiempos de espera en cola en aquellos casos en los cuales el paquete se genera pero no hay canal disponible para su transmisión. Junto con los parámetros de QoS anteriores se ha considerado un parámetro de QoS adicional cuya finalidad es la de cuantificar el nivel de satisfacción de los usuarios del sistema. El criterio empleado para definir este parámetro depende del servicio considerado. Para los servicios de navegación web y correo electrónico se considera que la transmisión de una página web o un correo electrónico resulta satisfactoria para el usuario cuando se completa en un periodo tiempo no superior a 4 segundos, según se indica en las especificaciones [154]. Para el servicio de vídeo H.263 en tiempo real, se considera que la transmisión de una trama de vídeo resulta satisfactoria para el usuario cuando finaliza antes de que se genere la siguiente trama de vídeo. El nivel de satisfacción de los usuarios de un determinado servicio se define como el porcentaje de ocasiones en las que la transmisión de una página web, correo electrónico o trama de vídeo resulta satisfactoria según el criterio correspondiente a cada servicio. Los posibles tiempos de espera en cola también se tienen en cuenta a la hora de calcular el valor de este parámetro.

En el presente trabajo se han considerado distintos tipos de servicios. Concretamente, los servicios considerados han sido de navegación web y correo electrónico así como de vídeo H.263 en tiempo real a diferentes tasas binarias medias (32, 64 y 256 kbps). El rendimiento de los algoritmos propuestos se ha evaluado bajo diferentes escenarios de carga en el sistema (baja/alta) a fin de poder mostrar claramente ciertos efectos que tienen lugar bajo determinadas condiciones de carga. La tabla 5.1 muestra el número de usuarios por celda de cada servicio para cada uno de los escenarios considerados.

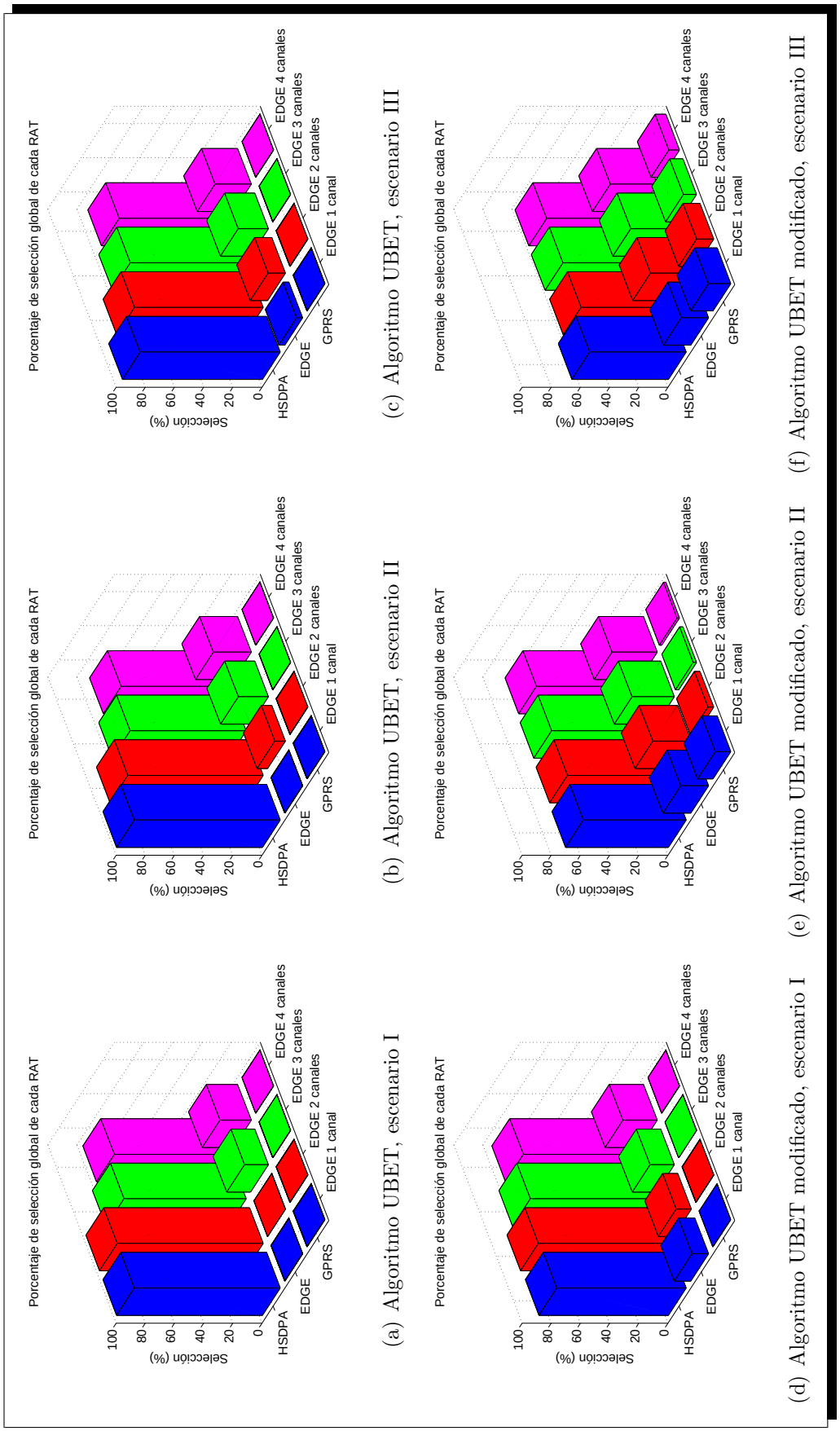
Escenario	Número de usuarios					Total
	Navegación web	Correo electrónico	Vídeo 32 kbps	Vídeo 64 kbps	Vídeo 256 kbps	
I	2	2	4	3	1	12
II	3	3	6	4	2	18
III	4	4	8	6	2	24

Tabla 5.1: Escenarios de evaluación.

5.2. Rendimiento del algoritmo UBET

El algoritmo UBET, según se discutió en el capítulo 3, fue diseñado con el objetivo de obtener un elevado nivel global de QoS en el sistema. Para alcanzar tal fin, el algoritmo UBET asigna los usuarios del sistema a aquella RAT que, en cada momento, ofrece el valor de *throughput* estimado más elevado. Teniendo en cuenta esto, es fácil entender los resultados mostrados en la figura 5.1, en la cual se muestra el porcentaje de selección global de cada RAT del sistema para el algoritmo UBET considerando los escenarios I, II y III (figuras 5.1a, 5.1b y 5.1c respectivamente). Las figuras muestran los valores globales, aunque los valores obtenidos para cada servicio de manera individual son en general muy próximos a la media global mostrada en dichas figuras. Cada figura muestra los resultados obtenidos cuando EDGE funciona en modo tanto *monocanal* como *multicanal*, es decir, cuando a cada usuario de EDGE se le asigna, respectivamente, un único canal (situación etiquetada como "*EDGE 1 canal*") o varios canales de manera simultánea (concretamente hasta 2, 3 y 4 canales por usuario). Para GPRS y HSDPA se considera en todos los casos operación monocanal. Según puede apreciarse en la figura 5.1a para el escenario I, cuando todas las RATs operan en modo monocanal, todos los usuarios del sistema son siempre asignados a HSDPA. Este comportamiento se debe a la considerable diferencia existente entre las velocidades de transmisión ofrecidas por HSDPA y las ofrecidas por el resto de RATs cuando todas operan en modo monocanal. Baste recordar que las velocidades máximas teóricas de GPRS, EDGE y HSDPA operando en modo monocanal son, respectivamente, de 21.4 kbps, 59.2 kbps y 720 kbps (ver apéndice A). Esta importante diferencia de prestaciones en términos de régimen binario da lugar a que el valor de la función de utilidad U_1 del algoritmo UBET sea considerablemente superior para HSDPA. Incluso cuando el número de usuarios en HSDPA (12 en total para el escenario I) es hasta tres veces mayor que el número total de canales disponibles (ver tabla 4.6 en el capítulo 4), es decir, $U_2 = 1/3$, HSDPA resulta todavía preferible para el algoritmo UBET ya que, tal y como se ha dicho anteriormente, las velocidades binarias de HSDPA son notablemente superiores a las del resto de RATs del sistema cuando todas ellas operan en modo monocanal. Estas observaciones también resultan aplicables a los escenarios II y III (con la salvedad de que en este último caso se observa un tímido porcentaje de utilización de EDGE, del 4 %, producido por un valor $U_2 \approx 1/6$ para HSDPA).

Con la finalidad de que otra RAT del sistema pueda ofrecer unas prestaciones binarias comparables a las ofrecidas por HSDPA, hacemos que la siguiente RAT con mejores prestaciones binarias del sistema, es decir, EDGE, opere en modo multicanal. Cuando EDGE opera en modo multicanal se observa que el porcentaje de selección de EDGE aumenta. Esta tendencia se acentúa a medida que aumenta el número de canales simultáneamente asignados a cada usuario en EDGE. Este comportamiento indica que cuando las RATs del sistema poseen unas capacidades binarias considerablemente diferentes, el algoritmo UBET selecciona de forma masiva aquella con mejores prestaciones. En cambio, cuando las capacidades de transmisión de las RATs disponibles resultan comparables, los usuarios del sistema son distribuidos entre ellas. A fin de poder verificar esta última observación, se simuló una red heterogénea integrada únicamente



(a) Algoritmo UBET, escenario I (b) Algoritmo UBET, escenario II (c) Algoritmo UBET, escenario III

(d) Algoritmo UBET modificado, escenario I (e) Algoritmo UBET modificado, escenario II (f) Algoritmo UBET modificado, escenario III

Figura 5.1: Porcentaje de selección global de cada RAT para el algoritmo UBET.

por GPRS y EDGE, ambas operando en modo monocanal. Los resultados obtenidos mostraron que GPRS se seleccionó en un 33.82 % de las ocasiones mientras que EDGE fue seleccionada en un 66.18 % de las ocasiones, lo cual confirma la afirmación anterior.

Una vez analizado el comportamiento del algoritmo UBET ante diferentes situaciones, resulta interesante analizar cómo repercute dicho comportamiento sobre el nivel de QoS percibido por los usuarios del sistema. Si bien son varios los parámetros de QoS considerados y analizados en este trabajo, en este caso se presentarán los resultados correspondientes al nivel de satisfacción global de los usuarios, por resultar los más representativos. La tabla 5.2 muestra dichos resultados para cada uno de los escenarios y configuraciones considerados, junto con el número de cambios de RAT por usuario y por minuto. En primer lugar, es interesante señalar la obtención de elevados niveles de satisfacción global en el sistema por parte del algoritmo UBET. Para el escenario I el nivel de satisfacción es superior al 90 % en todos los casos. A medida que la carga del sistema aumenta, el nivel de satisfacción decrece, tal y como cabía esperar, debido a un aumento del nivel de interferencia en el sistema y a una reducción de la disponibilidad de recursos libres. Aun así, incluso en condiciones de elevada carga, como es el caso de los escenarios II y III, los niveles de satisfacción obtenidos están comprendidos entre el 80 % y el 90 % (excepto para el escenario III con EDGE operando en modo monocanal, para el cual se obtiene un nivel de satisfacción global del 71.64 %, que en cualquier caso puede considerarse aceptablemente alto). Estos resultados permiten afirmar que el algoritmo UBET se comporta de forma satisfactoria en la medida en que cumple el objetivo para el cual fue diseñado, es decir, el de ofrecer elevados niveles de QoS a los usuarios del sistema, los cuales experimentan un alto grado de satisfacción.

Por otra parte, según se señaló en los párrafos precedentes, el algoritmo UBET realiza cierta distribución de los usuarios entre las RATs del sistema cuando éstas poseen unas capacidades de transmisión equiparables. Este comportamiento resulta lógico si se tiene en cuenta que fue diseñado para seleccionar en todo momento aquella RAT con el mejor rendimiento estimado. Por lo tanto, a menos que otras RATs alternativas sean capaces de ofrecer unas prestaciones equiparables, el algoritmo UBET no debería dirigir los usuarios del sistema hacia dichas RATs, ya que ello supondría un descenso del nivel de QoS global. A medida que aumenta el número de canales simultáneamente asignados a los usuarios de EDGE y, por lo tanto, las capacidades de transmisión de EDGE y HSDPA resultan más próximas entre sí, el número de usuarios asignados a EDGE aumenta. El resultado general observado, según se aprecia en la tabla 5.2, es un aumento del nivel de satisfacción de los usuarios del sistema. Para los escenarios con cargas elevadas (escenarios II y III) esta tendencia resulta bastante clara. En ellos se observa que a medida que las prestaciones de EDGE y HSDPA se aproximan más entre sí, el algoritmo UBET predice unas prestaciones más semejantes entre ambas y asigna un mayor número de usuarios a EDGE, en lugar de dirigirlos hacia HSDPA. Como resultado, la carga de usuarios soportada por HSDPA decrece, lo cual redunda en la obtención de mejores prestaciones para los usuarios de HSDPA. Por ejemplo, el retardo normalizado en HSDPA para los escenarios II y III experimenta, respectivamente, una mejora relativa del 39.6 % y del 42.7 % cuando EDGE pasa de operar en modo monocanal a operar con 4 canales por usuario. Por otra parte, los usuarios que

Escenario	Parámetro	Modo de operación de EDGE			
		1 canal	2 canales	3 canales	4 canales
I	Satisfacción (%)	92.73	92.17	91.38	92.29
	Cambios/usuario/min.	0.00	0.00	1.11	1.37
II	Satisfacción (%)	83.09	84.02	87.99	89.51
	Cambios/usuario/min.	0.00	0.26	1.25	1.28
III	Satisfacción (%)	71.64	80.55	85.27	87.85
	Cambios/usuario/min.	0.28	0.78	1.06	1.18

Tabla 5.2: Satisfacción global de los usuarios con el algoritmo UBET.

son asignados a EDGE operando en modo multicanal experimentan unas prestaciones que, aunque ligeramente menores que las ofrecidas por HSDPA, tampoco resultan tan diferentes (el *throughput* medio global en los escenarios I, II y III desciende un 6.63 %, 9.72 % y 9.61 % para EDGE con 4 canales respecto a EDGE monocanal). A pesar de ello, a medida que las RATs del sistema ofrecen unas prestaciones más similares, la distribución de usuarios que realiza el algoritmo UBET se traduce en última instancia en una mejora del nivel de satisfación global de los usuarios del sistema, según muestra la tabla 5.2. El coste de esta mejora, según puede apreciarse, es un aumento del nivel de señalización debido a la mayor frecuencia de traspaso de usuarios entre las RATs del sistema. En el caso del escenario I también se observa un aumento del nivel de señalización del sistema, aunque en este caso el nivel de satisfacción de los usuarios permanece prácticamente constante. Esta circunstancia sugiere que en situaciones de carga moderada puede resultar desaconsejable operar las RATs con menores prestaciones en modo multicanal, ya que aunque la carga quede más repartida entre las RATs del sistema, ello no se traduce en un aumento del nivel de satisfacción de los usuarios, existendo no obstante un coste asociado al aumento de señalización. En situaciones de carga elevada, la estrategia anterior contribuye a mejorar el nivel de QoS ofrecido a los usuarios, según se observa para los escenarios II y III, con el coste ya citado.

En las líneas precedentes se ha analizado el comportamiento del algoritmo UBET en función de las capacidades de transmisión relativas de las diversas RATs disponibles en el sistema. A continuación se analizará el comportamiento del algoritmo UBET en función del nivel de carga soportado en el sistema. Comparando la figura 5.1a con las figuras 5.1b y 5.1c se observa, en líneas generales, un aumento del porcentaje de selección de EDGE a medida que la carga global del sistema aumenta. Según se vio anteriormente, el algoritmo UBET dirige preferiblemente los usuarios del sistema hacia la RAT con mayores capacidades de transmisión, que en el caso particular considerado en este trabajo es HSDPA. Este comportamiento se traduce en un elevado nivel de carga soportado por HSDPA cuando la carga global del sistema aumenta, lo cual da lugar a un deterioro de las prestaciones de HSDPA (por ejemplo, en las simulaciones se observó que el *throughput* medio en HSDPA experimenta un descenso relativo del 16.4 % al pasar del escenario I al escenario III, cuando EDGE opera con 4 canales por

usuario). Así pues, cuando la RAT con mejores prestaciones está fuertemente cargada y, como consecuencia de ello, las prestaciones de dicha RAT quedan deterioradas, cabe la posibilidad, según se discutió en el capítulo 3, de que otra RAT cuyas prestaciones máximas teóricas sean menores ofrezca, de hecho, unas prestaciones estimadas mejores. Esto es precisamente lo que ocurre en las figuras 5.1a, 5.1b y 5.1c. Por este motivo, a medida que aumenta el número total de usuarios del sistema, el porcentaje de selección de EDGE aumenta, lo cual implica que un mayor número de usuarios es traspasado desde HSDPA hacia EDGE contribuyendo así a evitar la saturación de HSDPA. Resulta interesante señalar que un mayor porcentaje de selección de EDGE (ver figura 5.1) no siempre lleva asociado necesariamente un mayor nivel de señalización (ver la tabla 5.2).

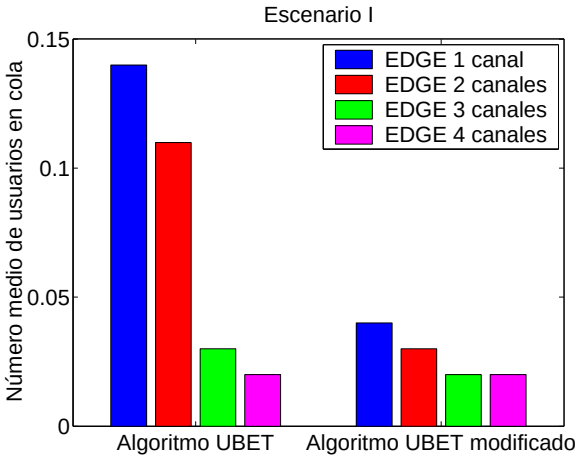
Hasta ahora se ha visto que el algoritmo UBET es capaz de ofrecer un alto grado de satisfacción a los usuarios del sistema, asignándolos a aquella RAT para la cual el nivel de prestaciones estimado es mayor. Esta estrategia implica, según se muestra en la figura 5.1, un elevado porcentaje de selección de la RAT con mejores prestaciones. Resulta interesante señalar que, como consecuencia del elevado porcentaje de selección de HSDPA, se observó durante las simulaciones que existe una cierta ocupación de la cola de espera de HSPDA aun cuando otras RATs del sistema disponían de canales libres. Por ejemplo, considerando en el escenario II el caso en el que todas las RATs operan en modo monocanal, el número medio de usuarios en cola en el sistema durante la simulación es de 1.8 usuarios. Puesto que el porcentaje de selección de HSDPA en este caso es del 100 %, esto implica que todos los usuarios en cola del sistema se hallan en la cola de HSDPA. Además, puesto que el porcentaje de selección de GPRS y EDGE en este caso es nulo, lo anterior implica que algunos usuarios del sistema permanencen en la cola de HSDPA esperando la liberación de un canal en HSDPA cuando todos los canales del resto de RATs del sistema permanecen ociosos. Es cierto que el resto de RATs del sistema posee unas capacidades de transmisión binarias notablemente menores y por este motivo el algoritmo UBET decide aun así asignar los usuarios a HSDPA. No obstante, algunos de los usuarios asignados a HSDPA deben esperar en la cola antes de recibir un canal de transmisión, mientras que si dichos usuarios fueran asignados a otra RAT con canales libres, la transmisión del paquete de datos comenzaría de inmediato y, a pesar de realizarse a una velocidad menor, tal vez podría terminar antes de que dicho usuario, permanenciendo en HSDPA, recibiera el canal y realizara la transmisión del paquete a una velocidad mayor. Ante esta reflexión surge la cuestión de si sería posible mejorar los niveles de QoS ofrecidos si en aquellos casos en los cuales el algoritmo UBET decide asignar un usuario a una RAT sin recursos libres se fuerza la asignación de dicho usuario hacia otra RAT que, aun ofreciendo prestaciones menores, dispone de recursos libres.

Para dar respuesta a la pregunta anterior se propone una versión modificada del algoritmo UBET en la cual antes de realizar la asignación de un usuario se verifica la disponibilidad de recursos en la RAT seleccionada. Si dispone de recursos libres, el usuario es asignado a dicha RAT, pero en caso contrario se verifica la disponibilidad de recursos en la RAT con el siguiente valor de utilidad más alto. El proceso se repite hasta encontrar una RAT con recursos libres. Si ninguna de las RATs dispone de recursos libres, el usuario finalmente es asignado a la RAT inicialmente seleccionada.

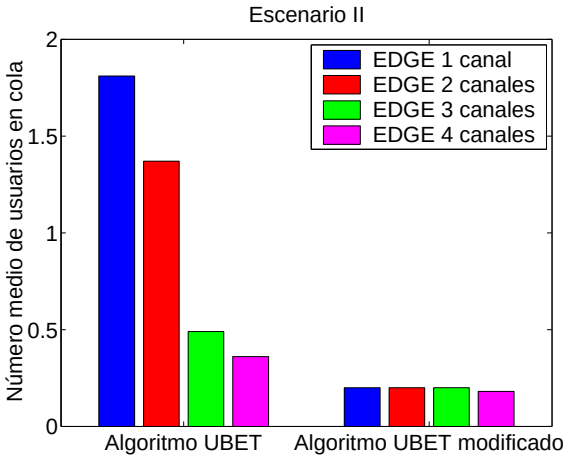
Las figuras 5.1d, 5.1e y 5.1f muestran el porcentaje de selección global de cada RAT del sistema para la versión modificada del algoritmo UBET. Las mismas observaciones generales que para el algoritmo UBET original resultan aplicables en este caso, es decir, el porcentaje de selección queda más *distribuido* entre las diversas RATs del sistema a medida que la carga del sistema aumenta o las prestaciones relativas de cada RAT son más semejantes entre sí. No obstante, es interesante observar que, para una misma situación, el algoritmo UBET modificado da lugar a un menor porcentaje de selección de HSDPA que el algoritmo UBET original y, por lo tanto, a un mayor porcentaje de selección del resto de RATs (especialmente llamativo resulta el porcentaje de selección de GPRS para el escenario III, que era nulo con el algoritmo UBET original en todos los escenarios). Este resultado confirma la observación anterior de que con el algoritmo UBET algunos usuarios permanecen en la cola de HSDPA mientras otras RATs disponen de recursos libres. Con la versión modificada de dicho algoritmo, los usuarios que hubieran acabado en la cola de HSDPA con el algoritmo UBET original son ahora asignados a otras RATs con recursos libres, dando como resultado un importante descenso en el número medio de usuarios en cola, según se aprecia en la figura 5.2 para los diversos escenarios considerados.

No obstante, todavía queda por verificar si la versión modificada del algoritmo UBET, gracias a la reducción observada en el número de usuarios en cola, es finalmente capaz de mejorar los niveles de QoS alcanzados por el algoritmo UBET original. Si bien es cierto que con la versión modificada del algoritmo se observan mejoras significativas de las prestaciones ofrecidas a los usuarios de HSDPA (gracias a una reducción del número de usuarios de dicha RAT), los resultados observados a nivel global en el sistema sufren un importante descenso. Por ejemplo, los valores medios de *throughput* y retardo normalizado en HSDPA llegan a experimentar mejoras de hasta un 7.0 % y un 61.3 % respectivamente con la versión modificada del algoritmo UBET, aunque los valores de dichos parámetros a nivel global en todo el sistema llegan a experimentar descensos de hasta un 24.6 % y un 41.3 % respecto al algoritmo UBET original. Estos resultados indican que el nivel de QoS global del sistema sufre un considerable descenso cuando se aplica la versión modificada del algoritmo UBET respecto a la versión original. Como resultado, el nivel de satisfacción de los usuarios disminuye, tal y como se muestra en la figura 5.3. El mejor funcionamiento del algoritmo UBET original respecto a la modificación propuesta queda también patente en la figura 5.4, donde se aprecia que no sólo ofrece mayores niveles de satisfacción a los usuarios sino que también lo consigue a un coste menor, es decir, introduciendo un menor nivel de señalización debida a los *handovers* verticales.

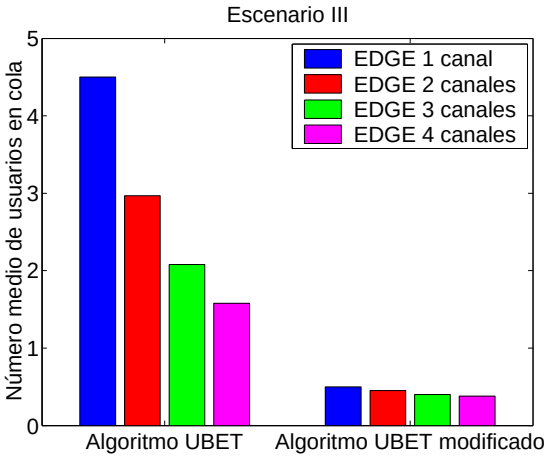
Los resultados anteriores confirman nuevamente la eficacia del algoritmo UBET a la hora de conseguir el objetivo propuesto y demuestran que, aunque las decisiones tomadas por dicho algoritmo impliquen en algunas ocasiones la existencia de usuarios en cola mientras una o varias de las RATs del sistema disponen de recursos libres, dichas decisiones redundan en última instancia en la obtención de mayores niveles de satisfacción para los usuarios del sistema respecto al caso en que éstos son traspasados a otras RATs con el objetivo de obtener un canal libre sin necesidad de tener que pasar primero por la cola de espera.



(a) Escenario I



(b) Escenario II



(c) Escenario III

Figura 5.2: Número medio de usuarios en cola para el algoritmo UBET.

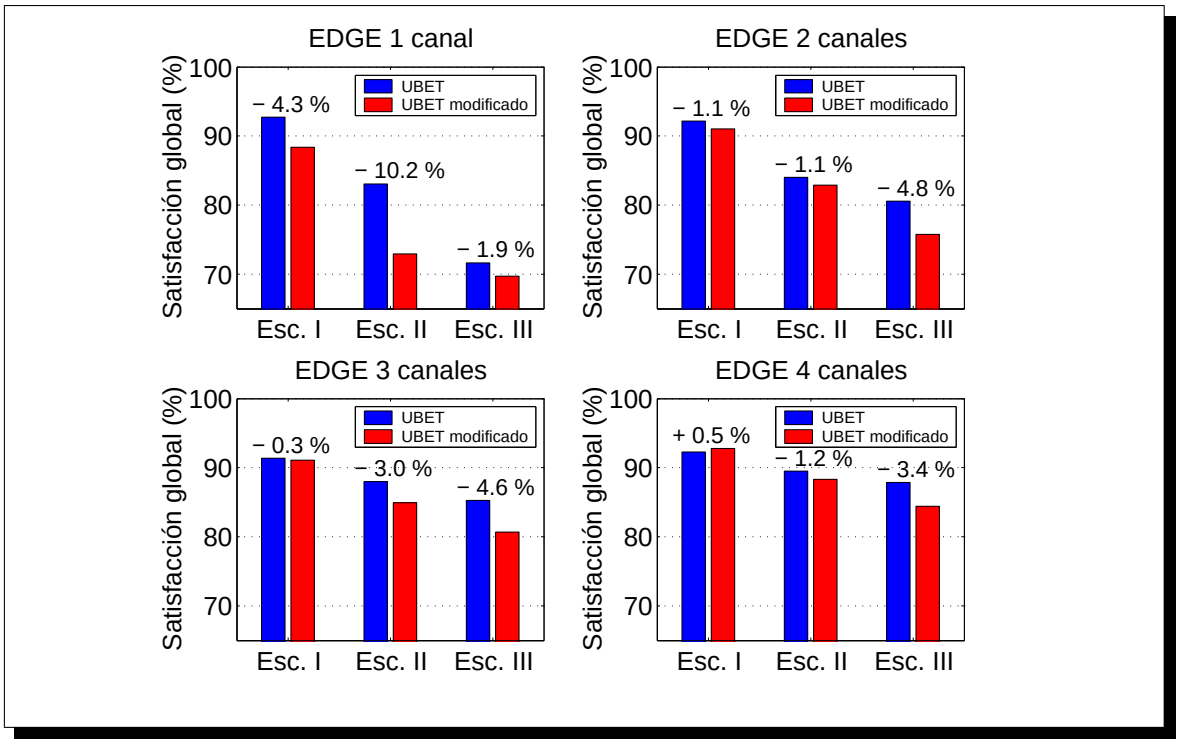


Figura 5.3: Satisfacción global de los usuarios para el algoritmo UBET.

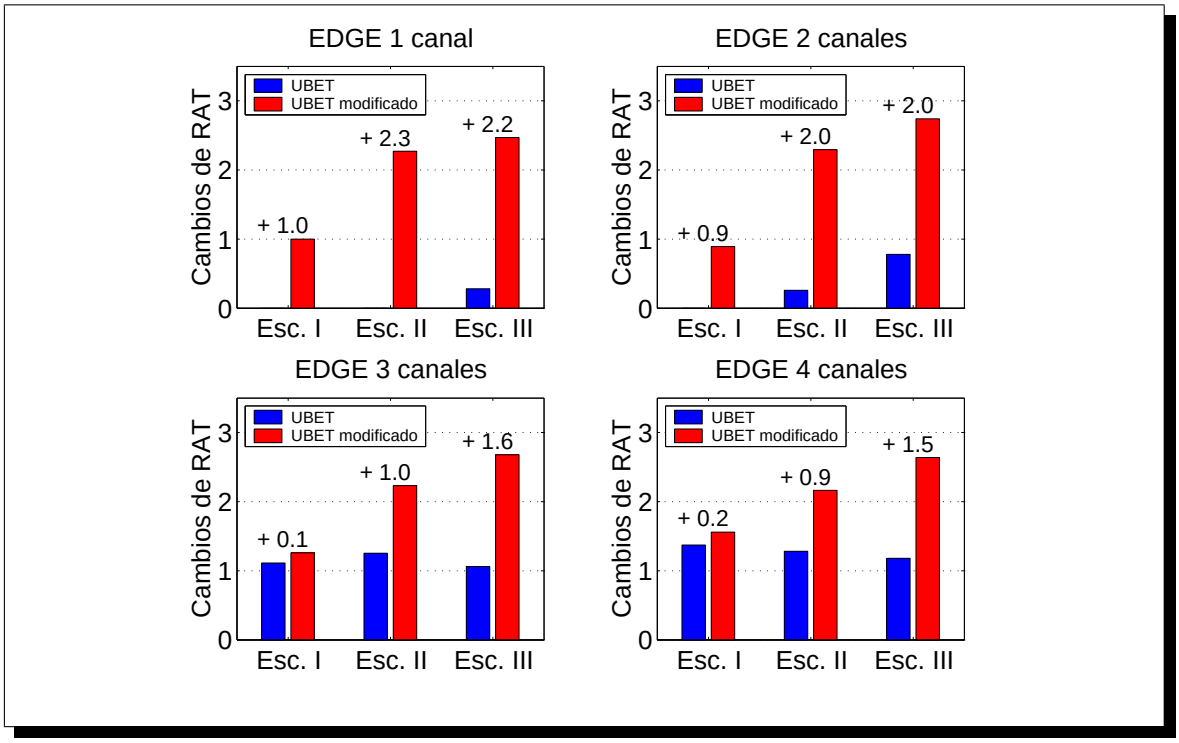


Figura 5.4: Número de cambios de RAT por usuario y por minuto para el algoritmo UBET.

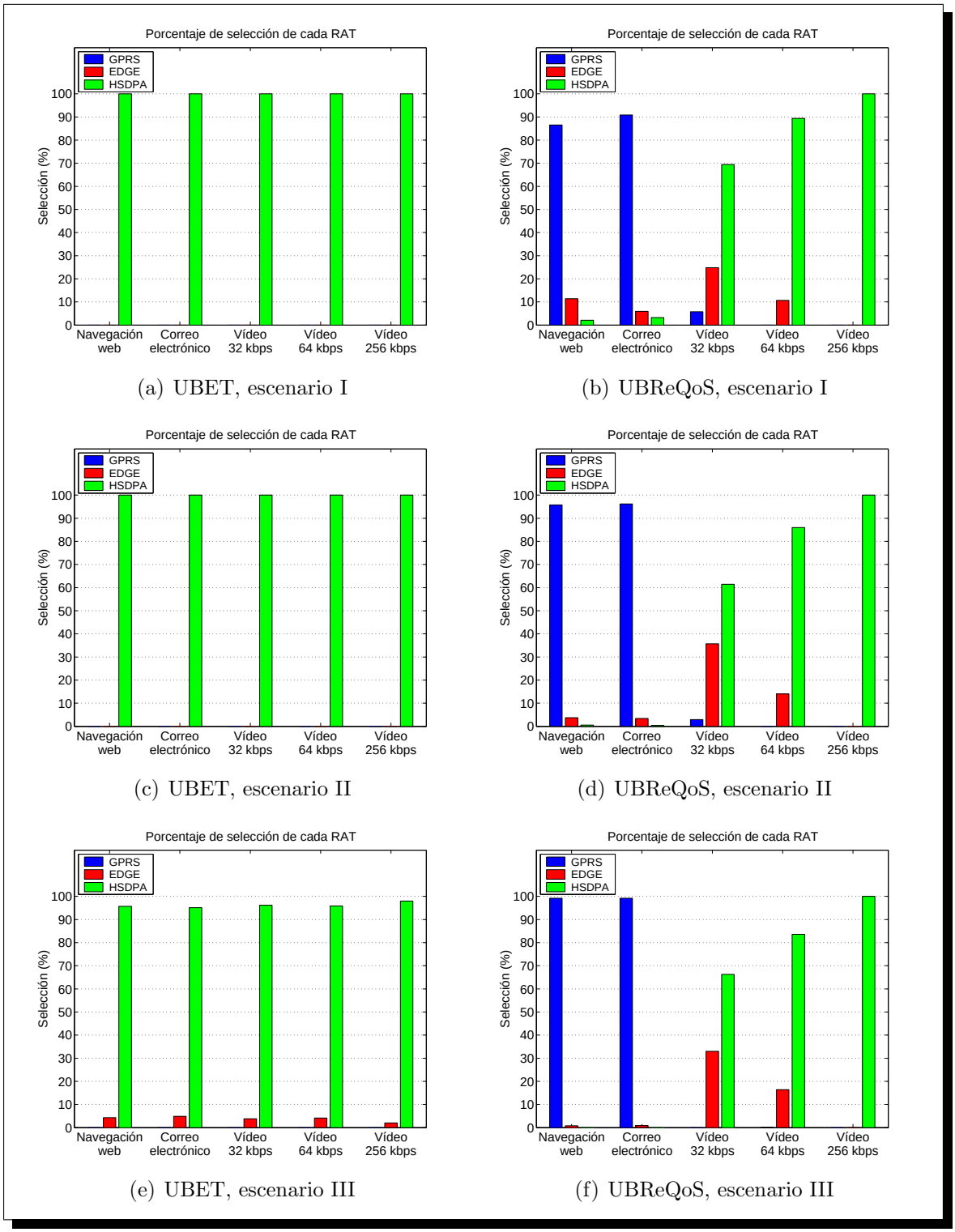
Existe un último aspecto importante del algoritmo UBET que, aunque puede deducirse de los resultados ya mostrados, no ha sido discutido de manera explícita y conviene, no obstante, remarcar. Según se dijo en el capítulo 3, una primera aproximación sencilla para la obtención de elevados niveles de QoS podría consistir en seleccionar siempre, de entre las RATs disponibles en la posición geográfica considerada, aquella que posea las mayores capacidades de transmisión en términos de tasa binaria, ignorando siempre las demás RATs. En el propio capítulo 3 se señaló que esta estrategia, aparte de suponer un uso claramente ineficiente de los recursos disponibles en el sistema, podría no proporcionar siempre los mejores resultados para los usuarios del sistema. Llegados a este punto se hace necesario verificar, así pues, si efectivamente la estrategia de asignar los usuarios del sistema siempre a la RAT con las mejores prestaciones puede verse superada, en términos de QoS, por otra estrategia que realice cierta distribución de usuarios entre las RATs disponibles y que, por lo tanto, no seleccione siempre la RAT con mejores prestaciones. En el caso particular considerado en este trabajo, la estrategia sencilla a la que se hace referencia consistiría en seleccionar siempre HSDPA, ya que el resto de RATs disponibles (GPRS y EDGE) poseen unas capacidades de transmisión más reducidas. La cuestión a resolver, así pues, es si otra estrategia que en algunas ocasiones seleccionara GPRS y/o EDGE podría llegar a ofrecer mejores niveles de QoS.

En la figura 5.1 se puede observar que, para los diversos escenarios considerados, cuando todas las RATs operan en modo monocanal se observa una utilización del 100 % de HSDPA por parte del algoritmo UBET (para el escenario III es del 96 % aunque, por simplicidad, consideraremos a efectos prácticos que es del 100 %). Esto indica que en este caso el algoritmo UBET se comporta precisamente como la estrategia consistente en seleccionar siempre la RAT con las mejores prestaciones. La tabla 5.2 indica que el nivel de QoS ofrecido por el algoritmo UBET cuando todas las RATs operan en modo monocanal o, lo que es equivalente, que el nivel de QoS ofrecido por la estrategia consistente en asignar los usuarios a la RAT con mejores prestaciones puede mejorarse cuando otras RATs alternativas operan en modo multicanal y se realiza cierta distribución de usuarios entre ellas. La estrategia anterior puede resultar satisfactoria cuando la carga del sistema es moderada, ya que en tal circunstancia el nivel de satisfacción de los usuarios permanece constante independientemente de que todos los usuarios del sistema se asignen a HSDPA (caso correspondiente a EDGE operando en modo monocanal) o se distribuyan entre HSDPA y EDGE operando en modo multicanal, según muestran los resultados obtenidos para el escenario I. No obstante, para cargas elevadas, la estrategia de asignar los usuarios a la RAT con mejores prestaciones no ofrece siempre el mejor rendimiento, ya que según queda patente en los resultados obtenidos para los escenarios II y III resulta posible obtener mayores niveles de satisfacción para los usuarios del sistema operando otras RATs en modo multicanal y realizando cierta distribución de usuarios entre ellas. En este caso el algoritmo UBET es un candidato adecuado para decidir cómo realizar dicha distribución, permitiendo así la obtención de un mayor nivel de satisfacción de los usuarios del sistema. Estos resultados demuestran, por lo tanto, que el algoritmo UBET constituye una mejor opción que la estrategia mencionada ya que, aunque ofrece un nivel de satisfacción similar para cargas bajas y moderadas, los usuarios experimentan un mayor nivel de satisfacción ante cargas altas.

5.3. Rendimiento del algoritmo UBReQoS

Los resultados mostrados en la sección 5.2 han demostrado que el criterio en que se basa el algoritmo UBET es capaz de obtener elevados niveles de satisfacción global asignando los usuarios del sistema a aquella RAT para la cual se presupone la obtención de las mejores prestaciones en términos de *throughput*. No obstante, tal y como se señaló en el capítulo 3, el algoritmo UBET decide las asignaciones sin tener en cuenta la posibilidad de que los requisitos de QoS de algunos usuarios con servicios poco exigentes podrían verse aceptablemente satisfechos aun cuando dichos usuarios fueran asignados a otras RATs con prestaciones más modestas, siempre y cuando resultaran suficientes para proveer el servicio de manera adecuada. Así pues se sugirió el algoritmo UBReQoS como alternativa al algoritmo UBET, el cual asigna los usuarios del sistema, no a la RAT con las mejores prestaciones estimadas, sino a aquella que resulte suficiente para satisfacer a los usuarios de los distintos servicios. La principal motivación de esta propuesta se basa en la idea de evitar que los usuarios de servicios con bajos requisitos de QoS, para los cuales puede resultar suficiente una red con prestaciones modestas, ocupen los recursos de las RATs con mejores prestaciones impidiendo de esta forma que otros usuarios de servicios con elevados requisitos de QoS y que, por tanto, necesitan imperativamente las RATs con mejores prestaciones, puedan recibir un servicio satisfactorio. Asignando los usuarios a una RAT cuyas prestaciones sean acordes al nivel de QoS requerido por el usuario, se podría evitar este tipo de situaciones, logrando que todos los usuarios del sistema reciban un servicio aceptablemente satisfactorio al mismo tiempo que se realiza un uso más eficiente de la totalidad de recursos radio disponibles en el sistema. Para tal fin se propone el algoritmo UBReQoS.

Como primer punto de comparación entre los algoritmos UBET y UBReQoS, la figura 5.5 muestra el porcentaje de selección de cada RAT, desglosado por tipo de servicio, para ambos algoritmos. En esta figura se muestran los resultados obtenidos para los diversos escenarios de carga considerados en este trabajo cuando todas las RATs operan en modo monocanal (los resultados obtenidos para otras configuraciones conducen a las mismas conclusiones). En primer lugar se observa claramente que el algoritmo UBET selecciona exclusivamente HSDPA para todos los servicios considerados, exceptuando el escenario III donde el porcentaje de selección global de EDGE es del 4 %. En el caso del algoritmo UBReQoS se observa, en cambio, que tanto para cargas bajas o moderadas (escenario I) como para cargas altas (escenarios II y III) los usuarios son distribuidos entre las diversas RATs disponibles de manera acorde a los requisitos de QoS de cada servicio (recordar que para la aplicación del algoritmo UBReQoS el nivel de QoS de cada servicio quedaba representado mediante su tasa binaria media a través del parámetro R_{nom}). Así, los usuarios de los servicios de navegación web y correo electrónico son asignados en la inmensa mayoría de ocasiones a GPRS, ya que dicha RAT ofrece unas prestaciones binarias suficientes para satisfacer los requisitos de QoS de estos servicios. En efecto, la tasa máxima de transmisión en GPRS en modo de operación monocanal es de 21.4 kbps, mientras que las tasas binarias requeridas por los servicios de navegación web y correo electrónico resultan ser de 5 y 12 kbps respectivamente (estos valores se han establecido según el procedimiento descrito en la sección



(a) UBET, escenario I

(b) UBReQoS, escenario I

(c) UBET, escenario II

(d) UBReQoS, escenario II

(e) UBET, escenario III

(f) UBReQoS, escenario III

Figura 5.5: Porcentaje de selección de cada RAT para los algoritmos UBET y UBReQoS.

3.2, considerando el valor de R_{nom} correspondiente al valor mediano del tamaño de los paquetes de datos generados por cada servicio). Por otra parte, los usuarios del servicio de vídeo a 256 kbps, el servicio más exigente de todos, son asignados a HSDPA en el 100 % de los casos. Ello es debido a que las prestaciones binarias ofrecidas por el resto de RATs resultan claramente insuficientes para los requisitos de dicho servicio. Para el servicio de vídeo a 64 kbps, la tasa máxima de transmisión que puede ofrecer EDGE (59.2 kbps) es inferior a la media requerida por dicho servicio, motivo por el cual se observa que los usuarios de este servicio son asignados habitualmente a HSDPA. En el caso particular del servicio de vídeo a 32 kbps se observa que la RAT seleccionada de forma más habitual es HSDPA, a pesar de que las prestaciones binarias de EDGE podrían resultar, en principio, suficientes para proveer dicho servicio. Sin embargo, las prestaciones reales ofrecidas por EDGE durante las simulaciones fueron apreciablemente menores que las máximas teóricas debido a que la propia configuración empleada para la plataforma de simulación utilizada en este trabajo (ver tabla 4.6) fomenta la existencia de considerables niveles de interferencia, aspecto que ya fue discutido en el capítulo 4 al final de la sección 4.3. Así pues, las prestaciones binarias ofrecidas por EDGE no llegan a ser lo suficientemente elevadas como para que los usuarios de vídeo a 32 kbps sean asignados habitualmente a EDGE. Esta circunstancia explica que los usuarios de vídeo a 32 kbps sean asignados a HSDPA en un mayor número de ocasiones. En cualquier caso, el aspecto de mayor relevancia que se observa en la figura 5.5 es que el algoritmo UBET asigna los usuarios de todos los servicios de manera habitual a la RAT con mejores prestaciones, mientras que el algoritmo UBReQoS los distribuye entre las RATs disponibles de manera acorde a los requisitos específicos de QoS del servicio utilizado y del nivel de prestaciones ofrecido por cada RAT.

La primera consecuencia derivada de la distribución de usuarios que realiza el algoritmo UBReQoS es un mayor grado de utilización de la totalidad de recursos disponibles en el sistema. Consecuentemente, se observa una significativa reducción en el grado de ocupación de las colas de espera del sistema respecto al algoritmo UBET. Así, mientras el número medio de usuarios en cola para los escenarios I, II y III con el algoritmo UBET era, respectivamente, igual a 0.14, 1.81 y 4.50 cuando todas las RATs operan en modo monocanal (ver figura 5.2), los resultados obtenidos con el algoritmo UBReQoS en estos mismos casos son, respectivamente, igual a 0.04, 0.38 y 1.49.

Sin embargo, la distribución de usuarios realizada por el algoritmo UBReQoS también implica que muchos de ellos dejan de ser asignados de forma habitual a HSDPA, cual era el caso del algoritmo UBET, para ser asignados con mayor frecuencia a otras RATs con menores capacidades de transmisión. Como resultado, la tabla 5.3 muestra (para operación monocanal) una significativa reducción respecto al algoritmo UBET del *throughput* obtenido por varios de los servicios considerados en este trabajo. El descenso más significativo se observa para los usuarios de navegación web y correo electrónico, los cuales han pasado de ser asignados habitualmente a HSDPA con el algoritmo UBET a ser dirigidos hacia GPRS en la mayoría de ocasiones con el algoritmo UBReQoS. Para los usuarios de vídeo a 32 y 64 kbps también se observa un descenso bastante significativo, ya que aunque sean asignados a HSDPA de manera más frecuente, ahora también son dirigidos en ciertas ocasiones hacia EDGE. A pesar

Servicio	Escenario I		Escenario II		Escenario III	
	UBET	UBReQoS	UBET	UBReQoS	UBET	UBReQoS
Navegación web	291.49	18.69	246.15	14.29	206.44	13.15
Correo electrónico	272.81	17.80	221.18	14.16	173.97	13.53
Vídeo a 32 kbps	237.28	104.34	203.21	78.21	173.48	81.92
Vídeo a 64 kbps	231.61	185.12	209.00	150.51	180.63	134.29
Vídeo a 256 kbps	233.27	248.14	210.45	240.83	185.60	237.59

Tabla 5.3: *Throughput* (kbps) para cada servicio con los algoritmos UBET y UBReQoS.

de la reducción observada en el *throughput* de estos servicios, es importante señalar que el valor medio final obtenido por todos ellos es superior al valor de R_{nom} utilizado en la aplicación del algoritmo UBReQoS. Así, por ejemplo, los usuarios de vídeo a 32 kbps experimentan un *throughput* bastante superior a 32 kbps en todos los escenarios. Esto significa que aunque el nivel de prestaciones ofrecido a estos servicios se haya visto reducido de manera importante con el algoritmo UBReQoS, sigue resultando, no obstante, superior al requerido por todos ellos para ser provistos de manera adecuada.

En la tabla 5.3 también se aprecia que el algoritmo UBReQoS no sólo provee los servicios anteriores con un nivel de prestaciones adecuado sino que, además, el nivel de prestaciones obtenido por los usuarios de vídeo a 256 kbps experimenta una mejora bastante significativa respecto al algoritmo UBET. Concretamente, las mejoras relativas observadas en el *throughput* de los usuarios de vídeo a 256 kbps para el algoritmo UBReQoS respecto al algoritmo UBET son del 6.37 %, 14.44 % y 28.01 % para los escenarios I, II y III respectivamente. Gracias a la distribución de tráfico realizada por el algoritmo UBReQoS, el nivel de carga soportado por HSDPA desciende respecto al algoritmo UBET. Como consecuencia, las prestaciones ofrecidas a los usuarios de vídeo a 256 kbps, que son asignados en el 100 % de los casos a HSDPA, experimentan una mejora significativa. En el caso particular de los usuarios de vídeo a 256 kbps, y a diferencia del resto de servicios considerados, se observa que a pesar de las mejoras obtenidas con el algoritmo UBReQoS respecto al algoritmo UBET, el valor medio de *throughput* sigue sin ser superior al valor de R_{nom} correspondiente a dicho servicio, es decir, 256 kbps. No obstante, es necesario señalar una vez más que las prestaciones ofrecidas por HSDPA en términos de *throughput* (relativamente bajas en comparación con el máximo teórico, comparar las tablas 5.3 y A.3) son consecuencia de la configuración adoptada para la plataforma de simulación empleada, la cual, según se comentó anteriormente, fomenta la existencia de elevados niveles de interferencia, limitando así el nivel de prestaciones ofrecido por las RATs del sistema. En cualquier caso, es interesante señalar nuevamente cómo el valor de *throughput* obtenido para los usuarios de vídeo a 256 kbps mejora significativamente con el algoritmo UBReQoS respecto al algoritmo UBET mientras los demás servicios experimentan un nivel de QoS adecuado.

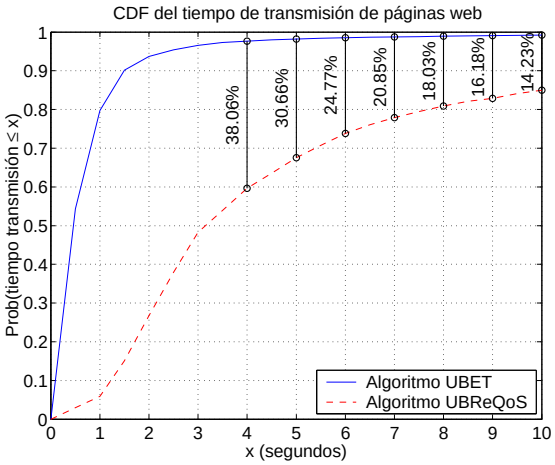
Los resultados mostrados en la tabla 5.3 indican que tanto el algoritmo UBET como el algoritmo UBReQoS son capaces de ofrecer un nivel de QoS superior al requerido por los diversos servicios considerados (a excepción del servicio de vídeo a 256 kbps por los motivos ya señalados). No obstante, el nivel de QoS ofrecido por el algoritmo UBReQoS resulta más acorde al nivel de QoS realmente requerido por cada servicio. En efecto, resulta paradójico observar por ejemplo cómo con el algoritmo UBET los usuarios de navegación web obtienen mejores prestaciones que los demás usuarios, incluso que los usuarios de vídeo. Gracias al reparto inteligente de usuarios realizado por el algoritmo UBReQoS, la provisión de servicios con requisitos de QoS bajos o moderados se realiza con unos niveles de QoS adecuados al mismo tiempo que se mejora el nivel de QoS ofrecido a los usuarios de aquellos servicios más exigentes. Estos resultados ponen de relieve la eficacia del algoritmo UBReQoS en la consecución del objetivo perseguido.

La tabla 5.4 muestra cómo repercuten los resultados anteriores sobre el nivel de satisfacción final experimentado por los usuarios de cada servicio. Considerando en primer lugar los resultados obtenidos para el escenario I, se observa que el nivel de satisfacción de los usuarios de navegación web y correo electrónico experimenta un significativo descenso con el algoritmo UBReQoS respecto al algoritmo UBET, concretamente, del 38.06 % y del 34.95 %. Cabe realizar dos observaciones al respecto.

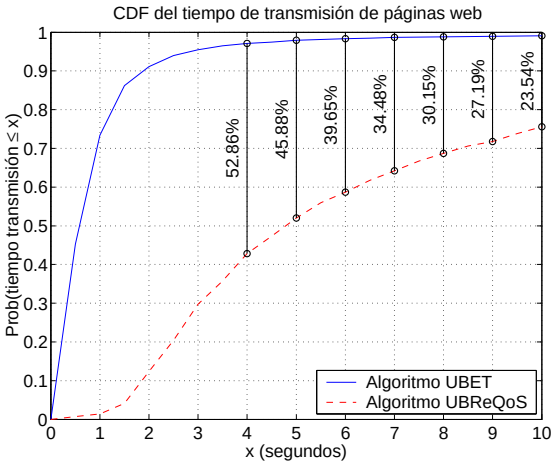
En primer lugar, conviene recordar que el nivel de satisfacción se define, para los usuarios de navegación web y correo electrónico, como el porcentaje de ocasiones en las que la transmisión de una página web o de un correo electrónico, respectivamente, se realiza en un intervalo de tiempo no superior a 4 segundos, tal y como se señala en [154]. No obstante, este valor es ciertamente relativo en la medida en que representa la percepción subjetiva de los usuarios del servicio. Esto quiere decir que si la transmisión de una página web o un correo electrónico se realiza en un intervalo de tiempo ligeramente superior, digamos de 5 ó 6 segundos, ello no implica necesariamente que esta transmisión resulte estrictamente no satisfactoria para el usuario. Así pues, aunque se haya seleccionado el valor de 4 segundos indicado en [154] para establecer un criterio de comparación común para todos los algoritmos y condiciones de evaluación considerados, es importante tener en cuenta la relatividad de los resultados obtenidos. Para ilustrar este hecho, la figura 5.6 muestra la CDF de los tiempos de transmisión de páginas web (para correo electrónico los resultados son similares). En la figura correspondiente al escenario I se observa que, para un valor de abscisa igual a 4 segundos, el valor de ordenada correspondiente al algoritmo UBReQoS es igual al 0.5964. Esto quiere decir que el 59.64 % de las páginas web se transmite en un intervalo de tiempo no superior a 4 segundos o, lo que es lo mismo, que si se considera un intervalo de tiempo de 4 segundos el nivel de satisfacción de los usuarios de navegación web es del 59.64 % con el algoritmo UBReQoS, un 38.06 % inferior al obtenido con el algoritmo UBET tal y como se muestra en la tabla 5.4. Si se toma como referencia un intervalo de tiempo de 5 ó 6 segundos, el nivel de satisfacción de los usuarios de navegación web con el algoritmo UBReQoS es entonces del 67.53 % y del 73.78 % respectivamente, lo cual supone una diferencia del 30.66 % y del 24.77 % respecto al algoritmo UBET, según se muestra en la figura 5.6a. A medida que se considera un intervalo de tiempo mayor, la diferencia del nivel de satisfacción de los usuarios de navegación web entre los algo-

Escenario	Servicio	Nivel de satisfacción		
		UBET	UBReQoS	Diferencia
I	Navegación web	97.70 %	59.64 %	–38.06 %
	Correo electrónico	92.80 %	57.85 %	–34.95 %
	Vídeo a 32 kbps	99.53 %	90.96 %	–8.57 %
	Vídeo a 64 kbps	96.05 %	90.30 %	–5.75 %
	Vídeo a 256 kbps	52.49 %	55.89 %	+3.40 %
	Global	92.73 %	85.92 %	–6.81 %
II	Navegación web	96.98 %	44.12 %	–52.86 %
	Correo electrónico	90.68 %	45.63 %	–45.05 %
	Vídeo a 32 kbps	95.20 %	92.38 %	–2.82 %
	Vídeo a 64 kbps	88.50 %	88.44 %	–0.06 %
	Vídeo a 256 kbps	42.58 %	54.62 %	+12.04 %
	Global	83.09 %	84.08 %	+0.99 %
III	Navegación web	94.72 %	34.16 %	–60.56 %
	Correo electrónico	88.63 %	31.50 %	–57.13 %
	Vídeo a 32 kbps	84.60 %	93.78 %	+9.18 %
	Vídeo a 64 kbps	72.19 %	85.03 %	+12.84 %
	Vídeo a 256 kbps	27.80 %	49.72 %	+21.92 %
	Global	71.64 %	84.18 %	+12.54 %

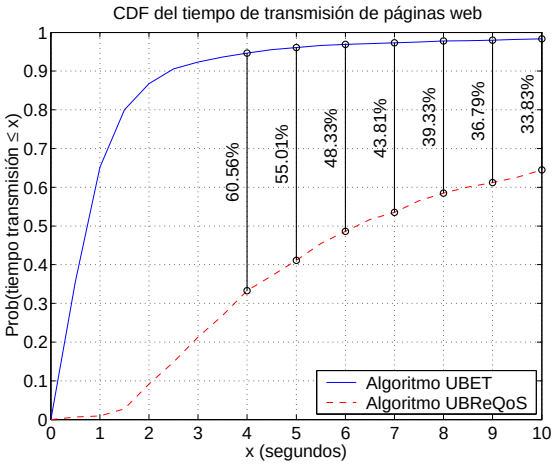
Tabla 5.4: Nivel de satisfacción de cada servicio con los algoritmos UBET y UBReQoS.



(a) Escenario I



(b) Escenario II



(c) Escenario III

Figura 5.6: CDF del tiempo de transmisión de páginas web.

ritmos UBET y UBReQoS es menor (lo mismo sucede para correo electrónico). Este hecho se observa también para los demás escenarios considerados (figuras 5.6b y 5.6c). En cualquier caso, los resultados obtenidos muestran que el nivel de satisfacción de los usuarios de navegación web y correo electrónico decrece con el algoritmo UBReQoS respecto al algoritmo UBET para todos los escenarios considerados.

En este punto cabe realizar una segunda observación. Los servicios de navegación web y correo electrónico son servicios que habitualmente son ofrecidos como servicios *best-effort*, es decir, servicios para los cuales el proveedor no garantiza unos niveles de QoS mínimos sino que simplemente se compromete a ofrecer el mejor nivel de QoS que resulte posible en cada momento (aun así la tabla 5.3 muestra que los valores medios de *throughput* experimentados por estos servicios son superiores a los valores de R_{nom} correspondientes). Lo anterior significa que los usuarios de este tipo de servicios, conscientes de esta circunstancia, no deberían considerar inaceptable un descenso en el nivel de prestaciones obtenido ya que además, generalmente, este tipo de servicios se facturan a un precio relativamente menor que otro tipo de servicios. Por otra parte, al proveedor le podría resultar interesante reducir el nivel de QoS ofrecido a los servicios menos exigentes (navegación web y correo electrónico en este caso) si con ello consigue incrementar el nivel de satisfacción de los usuarios de aquellos servicios con mayores requisitos de QoS, los cuales son a su vez los que se facturan a un precio más elevado. Este es, en efecto, el caso mostrado en la tabla 5.4. A cambio de reducir el nivel de satisfacción de los usuarios de navegación web y correo electrónico, se consigue incrementar el nivel de satisfacción de los usuarios más exigentes, es decir, los usuarios de vídeo a 256 kbps. Para cargas bajas o moderadas (escenario I) se obtiene una mejora apreciable del 3.40 % con el algoritmo UBReQoS respecto al algoritmo UBET, siendo especialmente significativa para cargas mayores, concretamente, del 12.40 % para el escenario II y del 21.92 % para el escenario III (el hecho de que el nivel de satisfacción de los usuarios de vídeo a 256 kbps no llegue al 56 % en el mejor de los casos se debe, según se dijo, a la configuración empleada en las simulaciones).

Respecto al nivel de satisfacción experimentado por los usuarios de los servicios de vídeo a 32 y 64 kbps, se observan diferentes comportamientos. Para el escenario I el nivel de satisfacción de los usuarios de estos servicios experimenta un descenso del 8.57 % y del 5.75 % respectivamente. No obstante, es importante señalar que aun así el nivel de satisfacción final de dichos usuarios es elevado, entorno al 90 % en ambos casos. Para el escenario II el nivel de satisfacción se mantiene aproximadamente constante, con tendencia a un ligero descenso. Nuevamente, el nivel de satisfacción sigue siendo elevado en ambos casos, entorno al 90 % aproximadamente. Finalmente, para el escenario III es interesante señalar un incremento del nivel de satisfacción de dichos usuarios respecto al algoritmo UBET, obteniéndose un valor no muy distinto al obtenido en el escenario II (en comparación con el algoritmo UBET), a pesar del incremento del nivel de carga.

También resulta importante señalar que, a medida que el nivel de carga del sistema aumenta, el nivel de satisfacción decrece en general para todos los servicios, según ilustra la tabla 5.4. No obstante, un examen detenido de los valores mostrados revela que con el algoritmo UBET un aumento de carga afecta especialmente al nivel de

satisfacción de los servicios con mayores requisitos de QoS (vídeo a 64 kbps y, especialmente, vídeo a 256 kbps), mientras el nivel de satisfacción de los usuarios menos exigentes desciende pero en menor medida que los anteriores. En cambio, con el algoritmo UBReQoS se observa que un aumento de carga perjudica en mayor medida a los servicios menos exigentes (navegación web y correo electrónico), mientras los servicios con mayores requisitos de QoS experimentan un descenso proporcionalmente menor en el nivel de satisfacción experimentado. Este comportamiento por parte del algoritmo UBReQoS resulta adecuado en la medida en que los usuarios que más pagan por sus servicios, ante un aumento de carga, deberían tener prioridad a la hora de recibir un servicio satisfactorio, aun a costa de reducir el nivel de satisfacción de los usuarios de servicios de tipo *best-effort*, los cuales se facturan a un precio menor.

De las discusiones realizadas en los párrafos precedentes se desprenden varias conclusiones importantes. En primer lugar, el algoritmo UBReQoS ofrece a los usuarios un nivel de QoS más acorde al requerido que el algoritmo UBET, lo cual implica un menor nivel de satisfacción para los servicios menos exigentes en términos de QoS (facturados a un precio menor) a cambio de poder ofrecer un mayor nivel de satisfacción a los usuarios de aquellos servicios con mayores requisitos de QoS (facturados a un precio mayor). Por otra parte, el algoritmo UBReQoS responde de manera más adecuada que el algoritmo UBET ante un aumento de carga, intentando preservar el nivel de satisfacción de los usuarios con mayores requisitos de QoS y que, por tanto, más pagan por sus servicios, a costa de reducir el de los usuarios de servicios menos exigentes, los cuales se facturan a un precio menor. Un tercer aspecto a resaltar es que el algoritmo UBReQoS distribuye los usuarios del sistema entre las RATs disponibles de manera más homogénea que el algoritmo UBET, lo cual implica un mayor grado de utilización de la totalidad de recursos disponibles en el sistema y, en definitiva, un uso más eficiente de los mismos. Esta última observación también resulta interesante desde el punto de vista económico ya que, según se señaló en el capítulo 1, al operador no le resulta provechoso incurrir en unos determinados costes de operación y mantenimiento para una red que no es utilizada y que, por lo tanto, no genera beneficios.

Al comienzo de la sección 3.2, cuando se discutió el objetivo perseguido por el algoritmo UBReQoS, se sugirió una primera aproximación como posible solución sencilla al problema planteado. Dicha solución consistía en asociar cada servicio ofrecido en el sistema a una RAT cuyas capacidades permitieran satisfacer los niveles de QoS requeridos por dicho servicio. Así, cada vez que un usuario solicitara un determinado servicio, éste le sería siempre ofrecido a través de la RAT correspondiente. Este esquema, si bien resulta atractivo por su sencillez, posee el inconveniente, según se indicó en la sección 3.2, de que un aumento en la demanda de un determinado servicio podría dar lugar a la saturación de la RAT asociada a dicho servicio, aun cuando otras RATs adecuadas para proveer dicho servicio dispusieran de recursos libres. Se señaló entonces la necesidad de diseñar otros principios más elaborados que permitieran satisfacer los distintos niveles de QoS requeridos haciendo un uso más eficiente de la totalidad de recursos radio disponibles en el sistema, proponiéndose para tal fin el algoritmo UBReQoS. Llegados a este punto resulta conveniente, así pues, comparar el rendimiento ofrecido por el algoritmo UBReQoS con el ofrecido por la aproximación anteriormente citada.

Esta solución estática ya fue propuesta y estudiada, entre otras, en la referencia [109]. Es interesante citar que dicho trabajo se engloba dentro del proyecto de investigación de ámbito europeo EVEREST (ver sección 2.2). En el caso particular de [109], se considera un escenario compuesto por GERAN y UTRAN sobre el cual se proveen dos servicios, concretamente, voz y navegación web. La estrategia básica propuesta en [109] consiste en asignar los usuarios de voz a GERAN y los usuarios de navegación web a UTRAN. Para la aplicación de esta idea en el contexto del presente trabajo, donde se considera un conjunto más amplio de servicios, las asignaciones se han decidido en función de los requisitos binarios de cada tipo de servicio y de las prestaciones ofrecidas por cada RAT. Así, los usuarios de navegación web y correo electrónico son asignados siempre a GPRS, los de vídeo a 32 kbps son servidos siempre a través de EDGE y, finalmente, los servicios de vídeo a 64 y 256 kbps se proveen a través de HSDPA. Esta aproximación será denominada en lo sucesivo *algoritmo de asignaciones fijas*.

En los valores mostrados en la figura 5.5 para el algoritmo UBReQoS se observa que, en general, cada servicio es asignado de forma habitual a una determinada RAT. No obstante, la provisión de un servicio dado no queda restringida al uso exclusivo de una única RAT, como es el caso del algoritmo de asignaciones fijas. El algoritmo UBReQoS ofrece más flexibilidad en este sentido ya que adapta de forma dinámica la distribución del tráfico de la red a las condiciones particulares de carga experimentadas en cada momento. Como resultado, el nivel de prestaciones ofrecido a los usuarios experimenta una importante mejora, según indica la tabla 5.5, donde se muestra el *throughput* medio para operación monocanal. Especialmente importante es el nivel de mejora experimentado por los usuarios de vídeo a 32 kbps. Con el algoritmo de asignaciones fijas los usuarios de dicho servicio son siempre asignados a EDGE, donde el *throughput* experimentado resulta insuficiente; con el algoritmo UBReQoS, en cambio, son asignados de manera más habitual a otra RAT capaz de ofrecer un nivel de prestaciones adecuado, es decir, HSDPA en este caso. Debido a este movimiento de usuarios de vídeo a 32 kbps hacia HSDPA por parte del algoritmo UBReQoS, el nivel de carga soportado por HSDPA aumenta respecto al algoritmo de asignaciones fijas. Como resultado, las prestaciones de los usuarios de vídeo a 64 kbps se ven deterioradas de forma significativa. Aun así, es importante señalar que las prestaciones finales obtenidas por dichos usuarios siguen siendo notablemente superiores a las requeridas. También resulta interesante observar que las prestaciones de los usuarios de vídeo a 256 kbps no sólo no se ven afectadas por el aumento de carga en HSDPA sino que, además, experimentan mejoras apreciables. Los resultados obtenidos para el retardo normalizado muestran un comportamiento análogo al observado para el *throughput*.

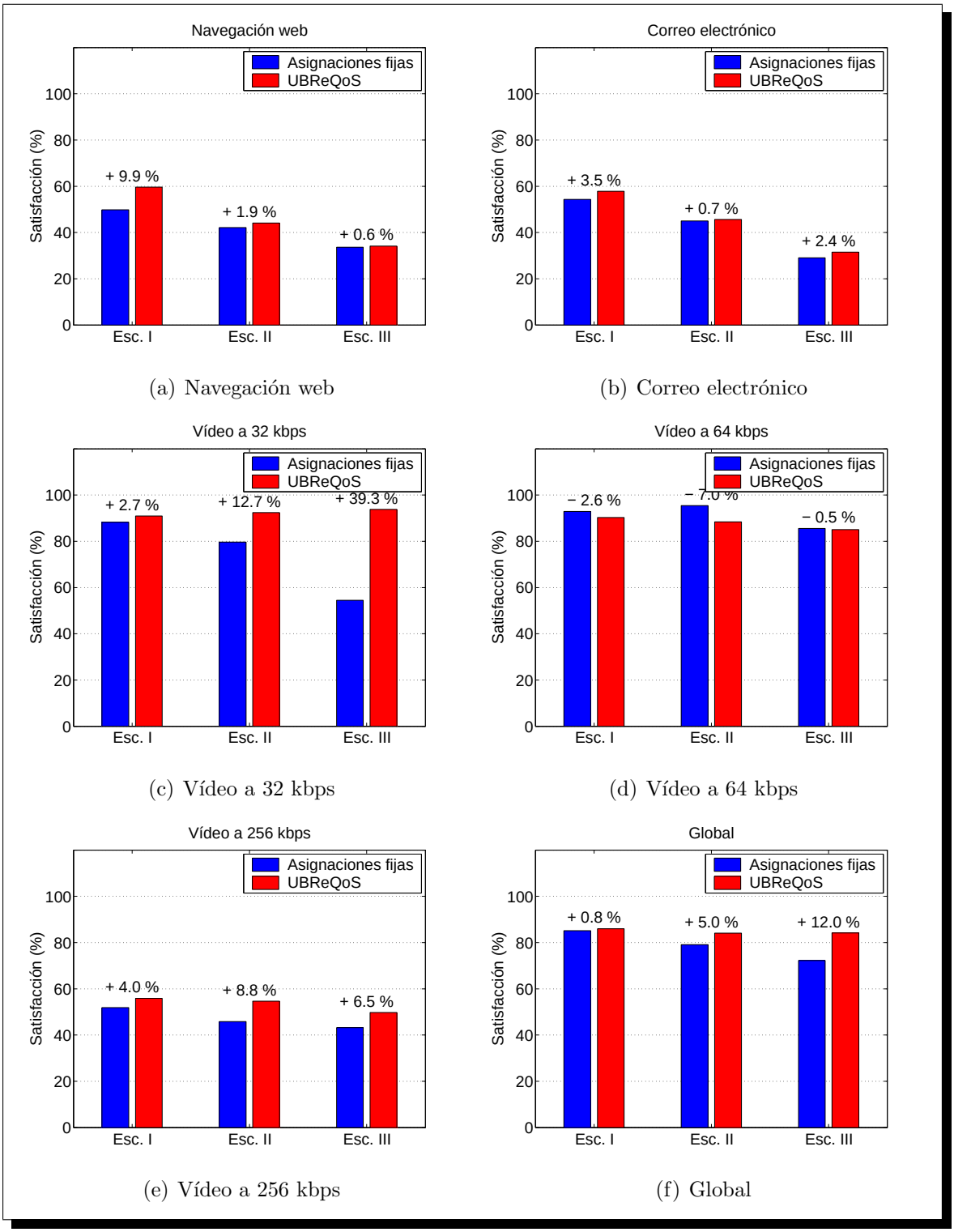
Como resultado de la mejor operación del algoritmo UBReQoS, los usuarios experimentan mayores niveles de satisfacción, llegándose a observar mejoras considerables. La figura 5.7 muestra los resultados obtenidos. Sólo los usuarios de vídeo a 64 kbps experimentan con el algoritmo UBReQoS menores niveles de satisfacción que con el algoritmo de asignaciones fijas, por los motivos explicados. En cualquier caso, el nivel de satisfacción de éstos sigue siendo elevado, entorno al 90 % aproximadamente. El coste de las mejoras introducidas por el algoritmo UBReQoS es el nivel de señalización generado por el traspaso de usuarios entre RATs, según muestra la tabla 5.6.

Escenario	Servicio	*Throughput* (kbps)		
		Asign. fijas	UBReQoS	Diferencia
I	Navegación web	15.64	18.69	+19.50 %
	Correo electrónico	15.41	17.80	+15.51 %
	Vídeo a 32 kbps	27.31	104.34	+282.06 %
	Vídeo a 64 kbps	227.91	185.12	–18.77 %
	Vídeo a 256 kbps	230.29	248.14	+7.75 %
II	Navegación web	13.31	14.29	+7.36 %
	Correo electrónico	14.00	14.16	+1.14 %
	Vídeo a 32 kbps	23.95	78.21	+226.56 %
	Vídeo a 64 kbps	219.35	150.51	–31.38 %
	Vídeo a 256 kbps	203.03	240.83	+18.62 %
III	Navegación web	13.12	13.15	+0.23 %
	Correo electrónico	12.71	13.53	+6.45 %
	Vídeo a 32 kbps	22.38	81.92	+266.04 %
	Vídeo a 64 kbps	216.53	134.29	–37.98 %
	Vídeo a 256 kbps	237.04	237.59	+0.23 %

Tabla 5.5: *Throughput* de cada servicio con los algoritmos UBReQoS y de asignaciones fijas.

Servicio	Número de cambios		
	Escenario I	Escenario II	Escenario III
Navegación web	0.25	0.29	0.06
Correo electrónico	0.05	0.04	0.01
Vídeo a 32 kbps	1.17	2.16	1.75
Vídeo a 64 kbps	0.40	1.15	1.17
Vídeo a 256 kbps	0.00	0.00	0.00
Global	0.54	1.03	0.89

Tabla 5.6: Número de cambios de RAT por minuto y usuario para el algoritmo UBReQoS.



(a) Navegación web (b) Correo electrónico

(c) Vídeo a 32 kbps (d) Vídeo a 64 kbps

(e) Vídeo a 256 kbps (f) Global

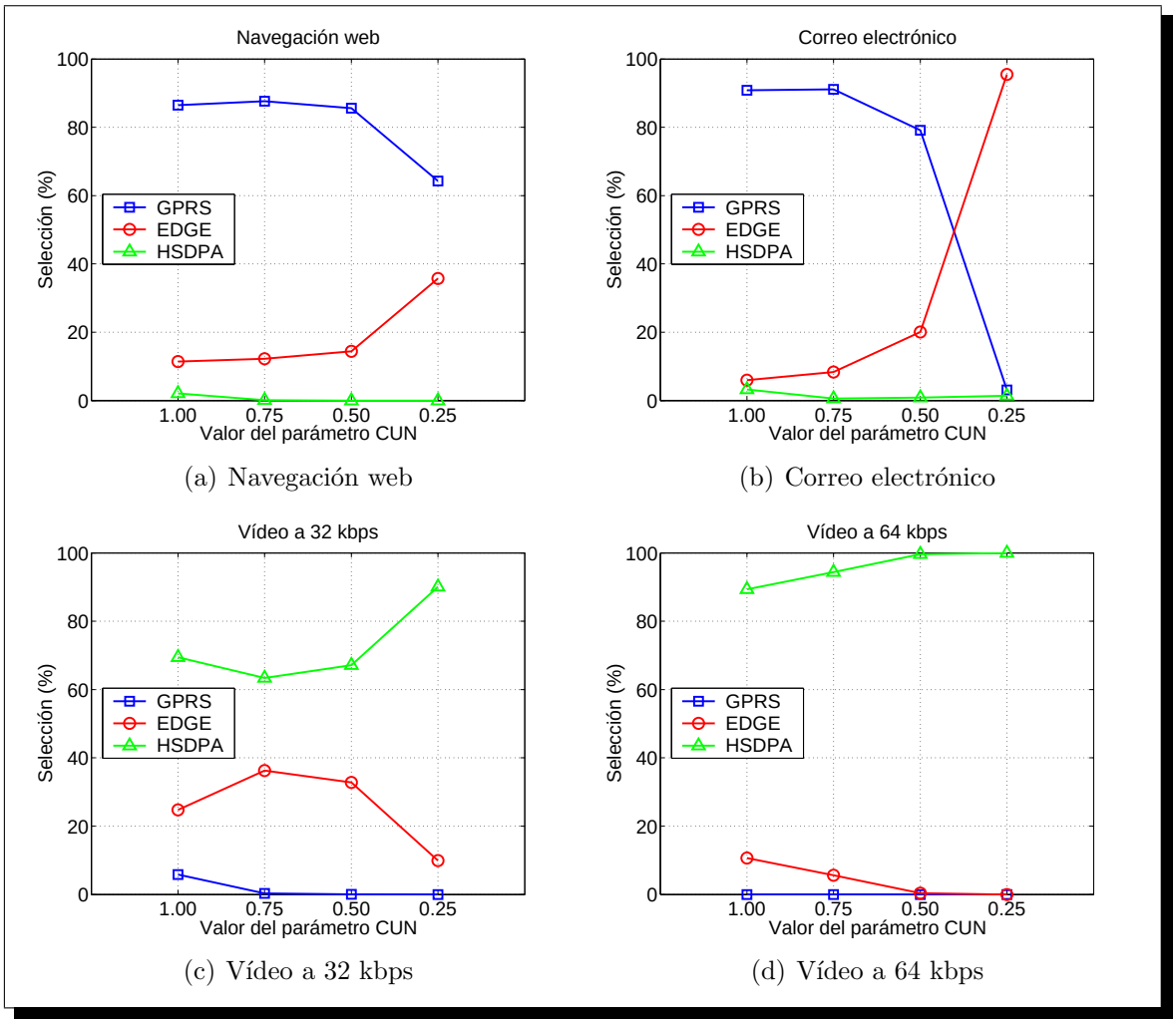
Figura 5.7: Nivel de satisfacción de cada servicio con los algoritmos UBReQoS y de asignaciones fijas.

Para concluir la evaluación del rendimiento del algoritmo UBReQoS, se analizará en último lugar cómo se comporta el algoritmo en función del valor asignado a sus parámetros de diseño. Según se mostró en la sección 3.2, en la definición del algoritmo UBReQoS se introduce un parámetro que se denominó *coeficiente de utilidad nominal* (CUN). Este parámetro representa el valor de utilidad que percibe el usuario cuando la tasa binaria esperada es igual a la tasa binaria media requerida por el servicio en cuestión (representada mediante el parámetro R_{nom}). Todos los resultados mostrados hasta el momento en la presente sección corresponden al caso en que CUN = 1.00. Este valor significa, dicho de manera un tanto tosca, que al usuario de un servicio que requiere una tasa binaria media igual a R_{nom} le resultará suficientemente útil una RAT capaz de ofrecer una tasa binaria igual o mayor que R_{nom}, mientras que en caso contrario dicha RAT no resultará suficientemente útil para el servicio considerado. Dicho de otro modo, un valor CUN = 1.00 implica que la tasa binaria mínima a partir de la cual una RAT determinada resulta suficientemente útil, es igual a la tasa R_{nom} requerida por el servicio considerado (ver figura 3.7). Por este motivo el valor CUN = 1.00 constituye, en principio, la elección más lógica, y ha sido este valor del parámetro CUN el que se ha utilizado a la hora de evaluar el rendimiento del algoritmo UBReQoS en las páginas precedentes. A continuación se analizará el efecto que provoca la aplicación de otros valores distintos para el parámetro CUN sobre el comportamiento observado del algoritmo UBReQoS.

Antes de mostrar los resultados obtenidos, merece la pena hacer un breve inciso para comentar la figura 3.7. Según se observa en dicha figura, a medida que se reduce el valor del parámetro CUN, aumenta el valor de la tasa binaria estimada mínima que debe ofrecer una determinada RAT para poder obtener el máximo valor de la función de utilidad u_1. Considérese que a la hora de tomar una decisión de selección RAT con el algoritmo UBReQoS, la tasa binaria estimada para las tres RATs consideradas en este trabajo, es decir GPRS, EDGE y HSDPA, es igual, por ejemplo, a 15 kbps, 40 kbps y 200 kbps respectivamente. Para el usuario de un servicio que requiera una tasa R_{nom} de 12 kbps cuya función de utilidad u_1 se calcule con CUN = 1.00, GPRS es capaz de ofrecer el máximo valor de u_1 según muestra la figura 3.7. En este caso la función de utilidad u_1 señalaría a GPRS como RAT candidata a ser seleccionada. Este comportamiento resulta, en principio, el más lógico, ya que para un servicio que requiera una tasa de 12 kbps, GPRS resulta suficiente si es capaz de ofrecer una tasa de 15 kbps. No obstante este comportamiento *lógico* sólo tiene lugar cuando CUN = 1.00 (por este motivo se seleccionó este valor a la hora de evaluar el rendimiento del algoritmo UBReQoS). Si se utiliza por ejemplo CUN = 0.5, la figura 3.7 indica que GPRS (cuya tasa binaria estimada es de 15 kbps) ya no ofrece el máximo valor de u_1. En este caso sólo EDGE (con 40 kbps) y HSDPA (con 200 kbps) ofrecen el máximo valor de u_1. Así pues, para CUN = 0.5 la función u_1 señalaría a EDGE como RAT candidata a ser seleccionada. Para un valor de CUN = 0.25, sólo la tasa binaria de 200 kbps estimada para HSDPA sería capaz de ofrecer el máximo valor de u_1. Resumiendo, para unos valores de CUN igual a 1.00, 0.50 y 0.25 las RATs seleccionadas, según la función u_1, serían, respectivamente, GPRS, EDGE y HSDPA en el ejemplo considerado. Así pues, se observa que a medida que decrece el valor del parámetro CUN existe una mayor probabilidad de seleccionar RATs con prestaciones binarias más elevadas.

Con la finalidad de confirmar la afirmación anterior, la figura 5.8 muestra, para cada servicio, el porcentaje de selección de cada RAT en función del valor del parámetro CUN, concretamente, para unos valores de CUN igual a 0.25, 0.50, 0.75 y 1.00. En el caso de navegación web y correo electrónico, que son asignados de manera habitual a GPRS, se aprecia claramente que una reducción del valor del parámetro CUN produce una reducción del porcentaje de selección de GPRS y, consecuentemente, un aumento del porcentaje de selección de EDGE. Para el servicio de vídeo a 32 kbps se observa un comportamiento análogo. Al pasar de CUN = 1.00 a CUN = 0.75 se observa que GPRS deja de ser seleccionada, aumentando el número de ocasiones en que se selecciona EDGE. Si el valor de CUN continúa decreciendo, el porcentaje de selección de EDGE disminuye y aumenta el de HSDPA. Finalmente, HSDPA pasa a ser la única RAT seleccionada para el servicio de vídeo a 64 kbps cuando decrece el valor del parámetro CUN. Para el servicio de vídeo a 256 kbps, no mostrado en la figura 5.8, se observó que para los valores de CUN considerados, HSDPA es seleccionada en el 100 % de los casos. Para que no fuera así sería necesario aplicar un valor de CUN superior a 1.00 (esta situación, no obstante, no ha sido evaluada). Los resultados mostrados en la figura 5.8 confirman, así pues, que una reducción del valor del parámetro CUN incrementa la frecuencia con que son seleccionadas las RATs con mayores prestaciones binarias.

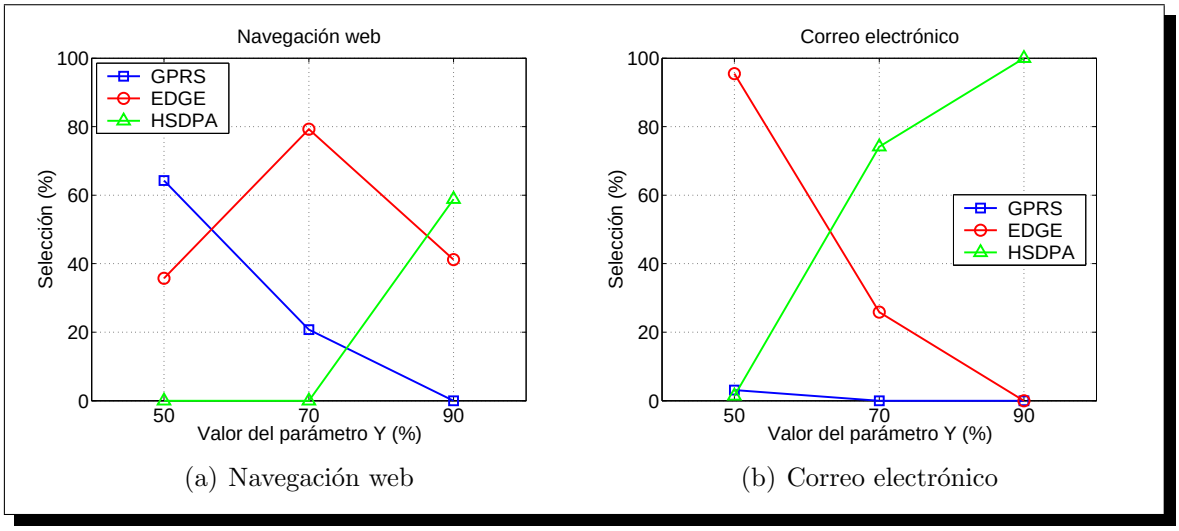
Existe un segundo parámetro que puede ser considerado como parámetro de diseño del algoritmo UBReQoS. En la sección 3.2 se propuso un método para determinar el valor del parámetro R_{nom} del algoritmo UBReQoS en el caso de servicios en tiempo no real como, por ejemplo, navegación web o correo electrónico. En dicho método el valor de R_{nom} se obtenía como el cociente entre un valor representativo del tamaño de los paquetes de datos generados por el servicio y un determinado intervalo de tiempo. El valor del primero se obtenía a partir de la CDF del tamaño de los paquetes de datos (ver figuras 3.9 y 3.10) y dependía del valor de ordenada considerado en dicha CDF (dicho valor de ordenada será referido en lo sucesivo como *parámetro Y*). Los resultados mostrados a lo largo de la presente sección corresponden al caso en que el valor del parámetro R_{nom} para los servicios de navegación web y correo electrónico se obtiene a partir del valor mediano del tamaño de los paquetes de datos generados o, lo que es equivalente, a partir del punto correspondiente al 50 % de dicha CDF, es decir, para Y = 50 %. Al final de la sección 3.2 se pudo ver que considerando otros valores mayores (Y = 70 % y 90 %), el tamaño de paquete obtenido era mayor y, por lo tanto, el valor de R_{nom} correspondiente era también mayor. Así pues, a medida que aumenta el valor del parámetro Y, aumenta el valor de la tasa binaria R_{nom} asociada a los servicios de navegación web y correo electrónico. Según los resultados ya mostrados en esta sección, si el valor de R_{nom} asociado a un servicio es mayor, existe un mayor porcentaje de selección de las RATs con mayores prestaciones binarias, siendo más elevado cuanto mayor es el valor de R_{nom}. Ante un aumento del valor del parámetro Y cabe esperar, por lo tanto, un incremento del porcentaje de selección de las RATs con mayores prestaciones binarias. La figura 5.9 muestra que en efecto es así. Para navegación web, al pasar de Y = 50 % a Y = 70 %, el uso de GPRS decrece y aumenta el de EDGE. Para Y = 90 %, el uso de GPRS y de EDGE decrece y aumenta el de HSDPA. Para correo electrónico, cuando el valor de Y crece, el porcentaje de selección de GPRS y EDGE desciende al mismo tiempo que aumenta el de HSDPA.



(a) Navegación web (b) Correo electrónico

(c) Vídeo a 32 kbps (d) Vídeo a 64 kbps

Figura 5.8: Porcentaje de selección de cada RAT en función del valor del parámetro CUN.



(a) Navegación web (b) Correo electrónico

Figura 5.9: Porcentaje de selección de cada RAT en función del valor del parámetro Y.

El análisis anterior muestra que una reducción del valor del parámetro CUN o un aumento del valor del parámetro Y da lugar a una mayor probabilidad de que el algoritmo UBReQoS seleccione RATs con prestaciones binarias más elevadas. Cuando esto sucede podría esperarse un aumento en el nivel de satisfacción experimentado por los usuarios. No obstante, los resultados obtenidos mostraron que esto no siempre es necesariamente así y, de hecho, en ocasiones puede llegar a suceder precisamente lo contrario. Hay que tener en cuenta que el nivel de satisfacción experimentado por los usuarios de cada servicio dependerá de la frecuencia con que sea seleccionada cada RAT para cada uno de ellos y esto, a su vez, podrá variar en diferente medida para cada servicio ante una misma variación del valor de los parámetros CUN o Y. De hecho, en las figuras 5.8 y 5.9 se observa que al variar el valor de estos parámetros en una determinada dirección, el porcentaje de selección de algunas RATs no varía de manera bien definida. Por ejemplo, en la figura 5.8c se observa que a medida que decrece el valor del parámetro CUN existe cierta tendencia de asignar dichos usuarios hacia RATs con mejores prestaciones, primero de GPRS a EDGE al pasar de CUN = 1.00 a CUN = 0.75, y luego de EDGE a HSDPA al seguir reduciendo el valor del parámetro CUN. Sin embargo, el porcentaje de selección de EDGE y HSDPA aumenta o decrece al reducir el valor del parámetro CUN. Algo similar sucede en la figura 5.9a para navegación web al variar el valor del parámetro Y. Como consecuencia de estos comportamientos, el nivel de satisfacción final experimentado por los usuarios de cada servicio es el resultado de complejos procesos de interacción entre todos ellos, lo cual no puede relacionarse de manera sencilla con el valor de los parámetros CUN e Y. Así pues, la única conclusión que puede extraerse de este análisis es que el valor empleado para los parámetros CUN e Y repercute directamente sobre las prestaciones ofrecidas por el algoritmo UBReQoS, si bien no resulta posible en este caso establecer una relación clara entre causa y efecto.

En cualquier caso, el análisis anterior no es en vano, ya que la conclusión obtenida sugiere la posibilidad de experimentar con diferentes combinaciones para los valores de los parámetros CUN e Y, con la intención de encontrar otra configuración distinta a la empleada a lo largo de esta sección que permita la obtención de mejores resultados. Con la finalidad de verificar esta posibilidad, se realizaron diversas pruebas considerando valores de CUN igual a 0.25, 0.50, 0.75 y 1.00, valores de Y igual al 50 %, 70 % y 90 % y, adicionalmente, EDGE operando en modo monocanal así como en modo multicanal con 2, 3 y 4 canales por usuario. En total son 48 las posibles combinaciones que pueden obtenerse para los valores indicados. Tras simular las 48 configuraciones señaladas se pudo comprobar que, en efecto, existen configuraciones distintas a la considerada en esta sección que mejoran los resultados obtenidos. Resulta bastante interesante señalar que dichas configuraciones llegan a abarcar todos los valores considerados para CUN, Y y el número de canales asignados a cada usuario en EDGE; no obstante, los mejores resultados obtenidos corresponden a valores bajos del parámetro CUN y a Y = 90 %. La tabla 5.7 compara el nivel de satisfacción obtenido con la configuración empleada en esta sección, etiquetada como configuración A, con el obtenido por las configuraciones para las cuales se observaron los mejores resultados tras realizar las pruebas anteriores. Los resultados mostrados en dicha tabla son de gran transcendencia para los objetivos de este trabajo, ya que muestran que con el algoritmo UBReQoS es posible servir a los usuarios del servicio más exigente (vídeo a 256 kbps en este

Configuración	A	B	C	D
Valor del parámetro CUN	1.00	0.25	0.25	0.50
Valor del parámetro Y	50 %	90 %	90 %	90 %
Número de canales en EDGE	1	1	3	4
Navegación web	59.64 %	94.01 %	93.33 %	96.57 %
Correo electrónico	57.85 %	93.51 %	94.80 %	93.06 %
Vídeo a 32 kbps	90.96 %	97.27 %	99.91 %	99.92 %
Vídeo a 64 kbps	90.30 %	94.96 %	94.96 %	97.47 %
Vídeo a 256 kbps	55.89 %	55.66 %	56.23 %	55.69 %
Global	85.92 %	90.81 %	92.13 %	93.09 %

Tabla 5.7: Nivel de satisfacción de cada servicio con el algoritmo UBReQoS para varias combinaciones de los parámetros de diseño del algoritmo (escenario I).

caso) con el máximo nivel de satisfacción observado (recordar la limitación impuesta por la configuración del simulador), al mismo tiempo que se incrementa el nivel de satisfacción de los demás servicios hasta alcanzar cotas superiores al 93 % en todos los casos, cercanas incluso al 100 % en algunos de ellos. Cuando el sistema opera con estas configuraciones se llega a observar una ocupación prácticamente nula de las colas de espera del sistema, una importante reducción de los valores de retardo normalizado y un aumento significativo del *throughput* experimentado por todos los servicios. Las mejores prestaciones de *throughput* quedan patentes, no sólo en el valor medio, sino también en el valor mínimo observado sobre un porcentaje bastante significativo de las medidas de *throughput* realizadas. La tabla 5.8 muestra, para cada servicio, el valor mínimo de *throughput* observado en el 90 % de las muestras tomadas cuando el sistema opera con la configuración B. Estos valores han sido obtenidos a partir de la CDF del *throughput* de cada servicio. Es interesante observar que el *throughput* mínimo observado sobre el 90 % de las muestras es superior al valor medio requerido por todos los servicios (excepto, una vez más, para vídeo a 256 kbps por las limitaciones que impone la configuración del simulador). En efecto, para Y = 90 %, que es el valor empleado para Y en esta configuración, el valor de R_{nom} correspondiente a los servicios de navegación web y correo electrónico obtenido según el procedimiento descrito en la sección 3.2 es de 31 kbps y 155 kbps, respectivamente. Para estos dos servicios el algoritmo UBReQoS ofrece un *throughput* mínimo de 49.60 kbps y 160.30 kbps en el 90 % de los casos. Para los servicios de vídeo a 32 y 64 kbps el *throughput* mínimo en el 90 % de las muestras tomadas también es apreciablemente mayor que la media requerida en ambos casos, no así para vídeo a 256 kbps por las causas ya señaladas. Estos resultados ponen de relieve que gracias a un ajuste adecuado de los parámetros de diseño del algoritmo UBReQoS es posible alcanzar un modo de operación óptimo que permite garantizar los diversos niveles de QoS requeridos en la red.

En contraste con los resultados anteriores cabe señalar que las configuraciones B, C y D dan lugar a un menor grado de utilización de la totalidad de recursos disponibles en el sistema respecto a la configuración A, ya que en estas configuraciones el porcentaje de selección de GPRS es nulo. Además, las mejoras introducidas acarrean un coste de señalización para las configuraciones C y D asociado a la mayor frecuencia de traspaso de usuarios entre RATs. No obstante, resulta bastante interesante señalar que el nivel de señalización introducido por la configuración B es notablemente menor que el correspondiente a la configuración A, tal y como puede apreciarse comparando los valores mostrados en la tabla 5.8 con los mostrados en la tabla 5.6 para el escenario I. Este hecho significa que la configuración B no sólo permite incrementar los niveles de satisfacción ofrecidos por la configuración A sino que además estas mejoras llevan asociadas una reducción del nivel de señalización asociada a la operación del algoritmo, lo cual afirma nuevamente la posibilidad de alcanzar una configuración óptima del algoritmo UBReQoS ajustando adecuadamente el valor de sus parámetros de diseño.

Para concluir la presente sección, la tabla 5.9 compara el nivel de satisfacción de cada servicio con los algoritmos UBET y UBReQoS. Para este último se muestran los resultados obtenidos tanto con la configuración A como con la configuración B. Para todos ellos se muestran los resultados correspondientes a los escenarios I y II. En los resultados mostrados para el escenario I, la configuración indicada como configuración B es la indicada en la tabla 5.7, que podría considerarse la más atractiva de las mostradas en dicha tabla ya que, aunque no alcanza los valores máximos de satisfacción, sí alcanza valores superiores al 93 % en todos los servicios (excepto vídeo a 256 kbps) introduciendo el mínimo nivel de señalización. Para el escenario II, la configuración B del algoritmo UBReQoS corresponde a los valores CUN = 0.25 e Y = 70 %, que de las observadas en modo de operación monocanal para el escenario II es la que mejores resultados ofrece. En ambos escenarios la configuración A es la que ha sido considerada a lo largo de esta sección para evaluar el rendimiento del algoritmo UBReQoS (CUN = 1.00 e Y = 50 %).

La tabla 5.9 muestra que las conclusiones derivadas de la comparación entre el algoritmo UBET y el algoritmo UBReQoS funcionando con una configuración arbitraria (configuración A) también son aplicables al caso en que el algoritmo UBReQoS funciona con una configuración óptima (configuración B). En efecto, en general se observa, para ambos escenarios de carga, que el algoritmo UBReQoS logra mejorar respecto al algoritmo UBET el nivel de satisfacción de los servicios más exigentes (vídeo a 256 kbps) a cambio de deteriorar el nivel de satisfacción de los servicios con menores requisitos en términos de QoS, es decir, navegación web y correo electrónico. También se observa que con el algoritmo UBReQoS, funcionando en ambas configuraciones, un aumento de carga produce un mayor grado de deterioro del nivel de satisfacción para los usuarios de servicios menos exigentes, manteniéndose aceptablemente estable para los servicios con requisitos de QoS moderados o altos. No obstante, el aspecto interesante mostrado en la tabla 5.9 es que configurando los parámetros de diseño del algoritmo UBReQoS hasta lograr una operación óptima del mismo (configuración B) se consigue minimizar el grado deterioro observado en el nivel de satisfacción de los usuarios de navegación web y correo electrónico. En efecto, comparando las configuraciones A y B del algorit-

Servicio	*Throughput* mínimo en el 90 % de las muestras	Cambios de RAT por minuto y por usuario
Navegación web	49.60 kbps	0.13
Correo electrónico	160.30 kbps	0.00
Vídeo a 32 kbps	47.65 kbps	0.13
Vídeo a 64 kbps	83.96 kbps	0.00
Vídeo a 256 kbps	125.44 kbps	0.00

Tabla 5.8: Rendimiento del algoritmo UBReQoS con la configuración B (escenario I).

Escenario	Servicio	Nivel de satisfacción		
		UBET	UBReQoS (A)	UBReQoS (B)
I	Navegación web	97.70 %	59.64 %	94.01 %
	Correo electrónico	92.80 %	57.85 %	93.51 %
	Vídeo a 32 kbps	99.53 %	90.96 %	97.27 %
	Vídeo a 64 kbps	96.05 %	90.30 %	94.96 %
	Vídeo a 256 kbps	52.49 %	55.89 %	55.56 %
	Global	92.73 %	85.92 %	90.81 %
II	Navegación web	96.98 %	44.12 %	62.74 %
	Correo electrónico	90.68 %	45.63 %	85.12 %
	Vídeo a 32 kbps	95.20 %	92.38 %	96.86 %
	Vídeo a 64 kbps	88.50 %	88.44 %	95.16 %
	Vídeo a 256 kbps	42.58 %	54.62 %	53.93 %
	Global	83.09 %	84.08 %	88.39 %

Tabla 5.9: Nivel de satisfacción de cada servicio con los algoritmos UBET y UBReQoS.

mo UBReQoS se observa que con una configuración óptima del algoritmo se mantiene el nivel de satisfacción de los usuarios de vídeo a 256 kbps al mismo tiempo que se observa un incremento del nivel de satisfacción para los demás servicios, especialmente significativo para navegación web y correo electrónico.

Un último aspecto interesante de la tabla 5.9 es que, según se dijo anteriormente, la configuración óptima del algoritmo UBReQoS (señalada como configuración B) para el escenario I no es la misma que para el escenario II. En efecto, mientras que dicha configuración para el escenario I corresponde a unos valores CUN = 0.25 e Y = 90 %, para el escenario II corresponde a unos valores CUN = 0.25 e Y = 70 %. Este hecho sugiere, de cara a la implementación del algoritmo en un sistema real, la posibilidad de variar la configuración del algoritmo UBReQoS de forma dinámica en función de diversos factores como, por ejemplo, el nivel instantáneo de carga del sistema, con la finalidad de obtener en todo momento un rendimiento óptimo del algoritmo UBReQoS.

5.4. Evaluación del método empleado para estimar la calidad del enlace radio

Los algoritmos propuestos en este trabajo se basan en el uso de funciones de utilidad que cuantifican los diversos factores y aspectos tenidos en cuenta por el criterio de selección aplicado. En la función de utilidad de ambos algoritmos existen ciertos términos cuyo valor se calcula a partir de la calidad experimentada en el enlace radio, expresada a través de la relación CIR/SIR. En la literatura se han propuesto diversos métodos que permiten realizar estimaciones precisas del valor de la relación CIR/SIR. Por ejemplo, para el caso particular de sistemas celulares TDMA, como sería el caso de GPRS o EDGE, en [192] se propone un método que permite obtener una estimación con un error de 0.3 dB en 200 ms, o con un error de 0.1 dB en 600 ms. En [193] se propone una versión modificada del método anterior capaz de obtener estimaciones con un error de 1 dB en 4.6 ms, o con un error de 0.3 dB en 60 ms. No obstante, este tipo de métodos suelen ser complejos y poseen un notable coste computacional. Para simplificar este proceso en la aplicación de los algoritmos propuestos, se sugirió un método para realizar una estimación sencilla de la relación CIR/SIR a partir de un determinado modelo de propagación y del número de canales interferentes activos (ver páginas 58-61). Evidentemente, la intención del método propuesto en este trabajo no consiste en obtener una estimación precisa, como es el caso de los métodos anteriormente citados, sino más bien en proporcionar un procedimiento sencillo que permita realizar una estimación del orden de magnitud a partir de la cual el algoritmo pueda tomar una decisión sin necesidad de incurrir en métodos complejos y computacionalmente costosos. La finalidad de esta sección es la de determinar cómo repercute sobre el rendimiento del sistema el hecho de basar las decisiones de selección de RAT sobre los valores de CIR/SIR obtenidos por el método propuesto en este trabajo. Para tal fin se comparará el rendimiento del sistema cuando se aplica el método propuesto con el obtenido cuando las decisiones de selección de RAT se toman a partir de los valores reales de relación CIR/SIR experimentados en el momento de tomar la decisión.

Con la finalidad de simplificar y abreviar el análisis llevado a cabo en esta sección, sólo se considerará el algoritmo UBReQoS ya que, de los dos algoritmos propuestos en este trabajo, éste es el que muestra, a la luz de los resultados obtenidos en las secciones 5.2 y 5.3, unas propiedades más atractivas. Además, los resultados mostrados en esta sección corresponden sólo a la configuración empleada en la sección 5.3, es decir, CUN = 1.00 e Y = 50 %, evaluada sobre el escenario I cuando todas las RATs operan en modo monocanal. Se realizaron pruebas con otras configuraciones y otras condiciones de simulación distintas y, aunque los valores numéricos obtenidos varían de un caso a otro, las conclusiones a las cuales se llega son las mismas en todos ellos.

Como primer punto de evaluación del método propuesto, se analizó la diferencia observada entre el valor de relación CIR/SIR real en el momento de tomar la decisión de selección de RAT (obtenido a partir de la posición real del usuario dentro de la celda) con el obtenido por el método propuesto en el momento de tomar la decisión (en el cual se consideran unas posiciones específicas para calcular los valores de portadora/señal e interferencia). La figura 5.10 muestra la CDF de la diferencia observada entre los valores obtenidos en ambos casos. Según se puede apreciar, sólo en un 33 % de las ocasiones el valor de CIR obtenido por ambos métodos para GPRS y EDGE difiere en menos de 3 dB. Para el valor de SIR obtenido para HSDPA, el valor del parámetro anterior es ligeramente superior, concretamente, del 47 %. Para márgenes de error más reducidos se observa un menor porcentaje de coincidencias. Por ejemplo, sólo en el 14 % de los casos existe un error inferior a 1 dB entre el valor obtenido por el método propuesto y el obtenido a partir de la posición real del móvil, según se aprecia en la figura 5.10. Estos resultados indican que los valores obtenidos por el método propuesto difieren considerablemente de los valores reales experimentados, tal y como cabía esperar. En efecto, el método propuesto se basa en unas determinadas posiciones estáticas del terminal móvil y estima el valor de CIR/SIR a partir del número de canales interferentes activos. En un sistema real, para un mismo número de canales interferentes activos, el valor exacto de CIR/SIR experimentado puede variar considerablemente en función de la posición real del terminal móvil respecto a su propia estación base y a las demás estaciones base interferentes, entre otros factores. Por este motivo se aprecian grandes diferencias entre el valor de CIR/SIR obtenido por el método propuesto y el obtenido a partir de la posición real del terminal móvil.

A pesar de las diferencias anteriores, cabe preguntarse cómo repercute sobre el valor de utilidad total de cada RAT el considerar uno u otro valor de CIR/SIR. Para dar respuesta a esta pregunta se evaluó, en cada una de las RATs del sistema, el porcentaje de ocasiones en las cuales el valor de utilidad correspondiente a la relación CIR/SIR obtenida con el método propuesto difería respecto al valor de utilidad correspondiente a la relación CIR/SIR real en una cantidad no superior al 10 % de éste último valor de utilidad. Para GPRS, EDGE y HSDPA se observó que tal condición se cumple en el 18 %, el 4 % y el 60 % de los casos, respectivamente. Esto significa que si el valor de utilidad que el algoritmo UBReQoS asigna a cada RAT se calculara a partir del valor real de la relación CIR/SIR, las utilidades asignadas a cada RAT diferirían notablemente de las asignadas cuando éstas se calculan a partir del valor de relación CIR/SIR obtenido por el método propuesto para tal fin en este trabajo.

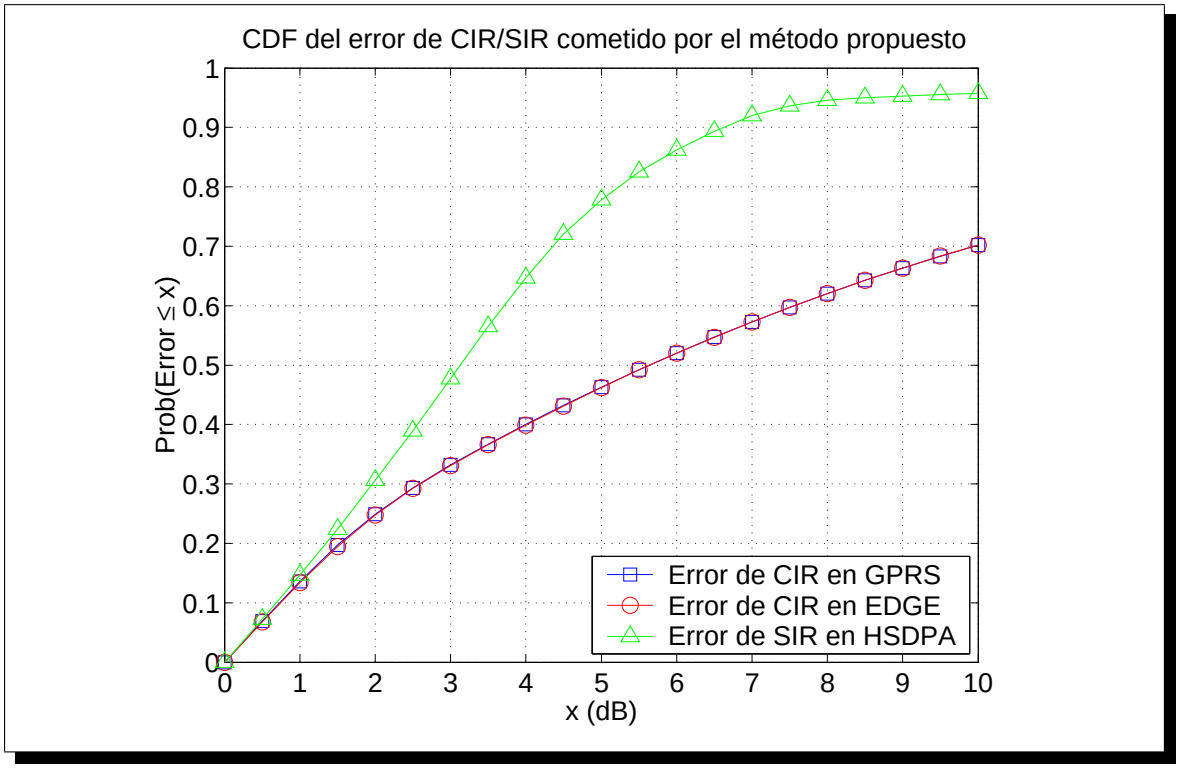


Figura 5.10: CDF del error de CIR/SIR cometido por el método propuesto.

A pesar de las diferencias observadas entre el valor de CIR/SIR real y el obtenido por el método propuesto, y de los diferentes valores de utilidad asignados a cada RAT dependiendo de cuál de los dos valores de CIR/SIR se utilice, todavía cabe preguntarse cómo repercuten estos hechos sobre la RAT que finalmente selecciona el algoritmo UBReQoS. En efecto, tiene sentido plantearse lo anterior porque, aunque el método propuesto introduzca un error notable en el valor de CIR/SIR estimado y esto afecte de forma significativa al valor de utilidad asignado a cada RAT, si dicho error afecta en igual medida al valor de utilidad asignado a todas las RATs del sistema, cabe esperar una cierta probabilidad de que la RAT finalmente seleccionada por el algoritmo sea la misma, tanto si se emplea el método propuesto como si se emplea el valor de CIR/SIR real. Para verificar esta posibilidad se contabilizó el número de ocasiones en las que la RAT seleccionada cuando se emplea el método propuesto es la misma que la que se seleccionaría si se empleara el valor real de CIR/SIR. Se observó que tal situación ocurre en un 76 % de los casos para las condiciones de evaluación consideradas (para el resto de condiciones de evaluación consideradas se obtuvieron resultados diferentes, aunque superiores al 60 % en todos ellos). Este resultado indica que existe un elevado porcentaje de ocasiones en las cuales la RAT seleccionada es la misma tanto si se considera el valor de CIR/SIR real como si se considera el obtenido por el método propuesto. No obstante, el porcentaje de ocasiones en que ambas decisiones difieren (entre el 30 % y el 40 % aproximadamente) resulta significativo, lo cual sugiere la posibilidad de que este porcentaje de ocasiones en que ambos métodos dan lugar a una decisión distinta pueda traducirse en una diferencia de prestaciones apreciable. Los resultados obtenidos en las simulaciones mostraron que, efectivamente, así es.

Servicio	Nivel de satisfacción		
	Método propuesto	Valor real	Diferencia
Navegación web	59.64 %	52.11 %	–7.53 %
Correo electrónico	57.85 %	57.40 %	–0.45 %
Vídeo a 32 kbps	90.96 %	93.43 %	+2.47 %
Vídeo a 64 kbps	90.30 %	76.79 %	-13.51 %
Vídeo a 256 kbps	55.89 %	49.14 %	-6.75 %
Global	85.92 %	80.85 %	-5.07 %

Tabla 5.10: Nivel de satisfacción de cada servicio.

La tabla 5.10 muestra el nivel de satisfacción de los distintos servicios considerados en este trabajo para el algoritmo UBReQoS con la configuración empleada en la sección 5.3, es decir, CUN = 1.00 e Y = 50 %. En la primera columna se muestra el nivel de satisfacción obtenido por los usuarios de los distintos servicios cuando el algoritmo UBReQoS decide la RAT óptima para cada usuario haciendo uso del valor CIR/SIR obtenido por el método propuesto en este trabajo (ver páginas 58-61). La segunda columna muestra el nivel de satisfacción obtenido para las mismas condiciones de evaluación cuando el mismo algoritmo decide la RAT óptima haciendo uso del valor de CIR/SIR real experimentado en el momento de tomar la decisión. Los resultados obtenidos indican que haciendo uso del método propuesto en este trabajo para obtener un valor de CIR/SIR se obtienen en general mejores resultados que haciendo uso del valor real de CIR/SIR experimentado en el momento de tomar la decisión.

El motivo de los resultados anteriores puede explicarse intuitivamente si se realiza una comparación entre el valor de CIR/SIR obtenido con el método propuesto y los posibles valores reales de CIR/SIR que puede experimentar un usuario en una posición dada. Para la tal fin, la figura 5.11 muestra el valor medio de la relación CIR que experimentaría un usuario de GPRS o de EDGE en función de su posición relativa dentro de la celda (la gráfica homóloga para HSDPA tendría un aspecto similar, aunque con valores distintos). Esta gráfica se ha obtenido haciendo uso del modelo de propagación implementado en la plataforma de simulación (modelo Okumura-Hata COST 231, ver sección 4.2.2) y considerando que existe un canal interferente en cada celda cocanal según la estructura celular empleada en el simulador (ver figura 4.2). El valor empleado para el radio de las celdas, potencia de transmisión, frecuencia portadora y demás parámetros relevantes son los mismos que se han considerado en el simulador (ver sección 4.3). Es importante señalar que el valor real que se obtendría en una determinada posición sería en general diferente del valor medio mostrado en la figura 5.11, ya que vendría determinado por otros diversos efectos de propagación, como el efecto multicamino (*multipath fading*) o el efecto sombra (*shadowing*), los cuales no quedan reflejados en dicha figura. No obstante, resulta suficiente considerar los valores medios para poder explicar intuitivamente los resultados mostrados en la tabla 5.10.

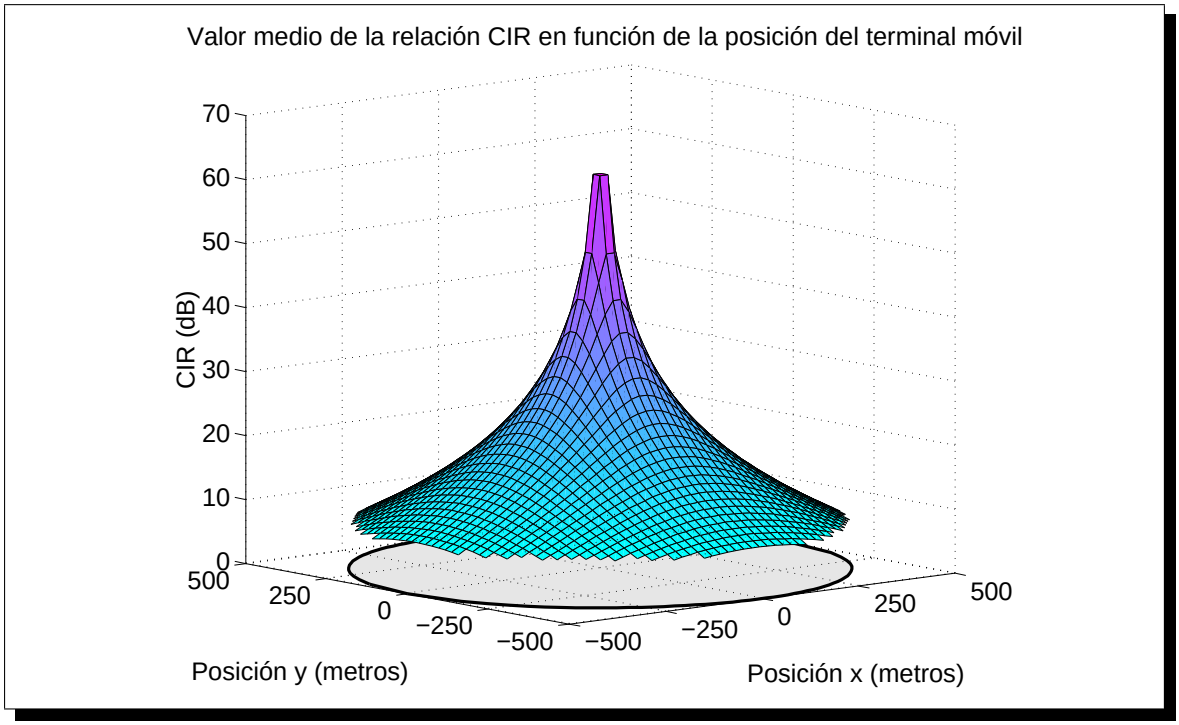


Figura 5.11: Valor medio de la relación CIR en función de la posición del terminal móvil.

En la figura 5.11 se observa que el valor medio de la relación CIR/SIR de un terminal móvil experimenta importantes variaciones, de hasta varias decenas de decibelios, en función de la posición en la que se encuentre dentro de la celda. En posiciones centrales, cerca de la estación base, se experimenta una buena calidad de transmisión radio mientras que a medida que el terminal móvil se aleja del centro y se dirige hacia el borde de la celda la calidad empeora. Si bajo las mismas condiciones de operación que las consideradas para obtener la figura 5.11 se calcula el valor de la relación CIR haciendo uso del método propuesto en este trabajo, el valor obtenido es de 8.57 dB. Comparando este valor con los mostrados en la figura 5.11 se observa que dicho valor es muy similar al obtenido en el peor caso, es decir, en el borde de la celda. Tras realizar varias pruebas se pudo verificar que esta afirmación sigue siendo cierta incluso para otras configuraciones celulares, potencias de transmisión, frecuencias portadoras, etcétera. Esta circunstancia indica, así pues, que el valor de CIR/SIR obtenido con el método propuesto en este trabajo corresponde aproximadamente al valor medio que se obtendría en el peor de los casos posibles, es decir, cuando el terminal móvil está en la posición más distante posible de su estación base. Conviene recordar en este punto que el método propuesto en este trabajo calcula el nivel de señal o portadora asumiendo que el terminal móvil se encuentra precisamente en dicha posición (ver figura 3.5). Según se señaló en su momento, el motivo por el cual se considera el peor caso es el de poder garantizar en cierta forma que el usuario va a recibir un servicio satisfactorio en la RAT seleccionada. Así, si el nivel de prestaciones ofrecido por la RAT seleccionada según el criterio del algoritmo UBReQoS resulta presumiblemente suficiente para satisfacer los requisitos de QoS del usuario en el peor de los casos, se ofrecen ciertas garantías de que los requisitos de QoS del usuario se verán satisfechos en la RAT seleccionada.

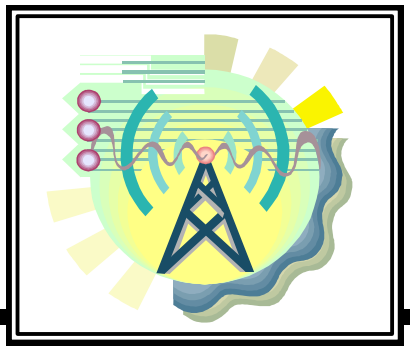
En cambio, cuando las decisiones de selección de RAT se toman a partir del valor de CIR/SIR real experimentado en el momento de tomar la decisión, no se puede garantizar un rendimiento mínimo al usuario porque las condiciones experimentadas en el momento de tomar la decisión de selección de RAT podrían variar significativamente durante el tiempo que permanece vigente dicha decisión. Hay que tener en cuenta que las decisiones de selección de RAT suelen permanecer vigentes durante un periodo de tiempo relativamente largo como, por ejemplo, el tiempo que dura una sesión del servicio, ya que no resulta muy sensato que un usuario pueda cambiar de RAT por ejemplo cada pocos segundos. Si la RAT óptima se decide a partir de unas condiciones específicas que tienen lugar en el momento de tomar la decisión pero que pueden variar de manera significativa durante el periodo de vigencia de dicha decisión, el resultado puede ser el mostrado en la tabla 5.10. Para ilustrar este hecho considérese el siguiente ejemplo. En el momento de tomar una decisión de selección de RAT para un usuario determinado, el valor de CIR real correspondiente a EDGE es tal que el nivel esperado de prestaciones de EDGE resulta suficiente para el servicio solicitado por dicho usuario. Consecuentemente, el criterio del algoritmo UBReQoS asignará dicho usuario a EDGE. No obstante, puede suceder que el usuario se mueva respecto a la posición en la cual se encontraba en el momento de tomarse la decisión y, como resultado, la calidad radio empeore notablemente. Incluso a una velocidad tan reducida como la considerada en este trabajo (3 km/h), la posición del usuario puede variar de manera significativa durante el tiempo que permance vigente la decisión de selección de RAT (por ejemplo, durante una sesión de 2 minutos se pueden recorrer 100 metros a 3 km/h). Si como consecuencia de este desplazamiento del usuario la calidad radio empeora lo suficiente, podría suceder que el nivel de prestaciones ofrecido por EDGE dejara de resultar suficiente para una provisión satisfactoria del servicio. En cambio, si la decisión se hubiera tomado a partir del valor de CIR correspondiente a la situación más desfavorable, cual es el caso del método propuesto en este trabajo, probablemente no se habría seleccionado EDGE sino que, en su lugar, podría haberse seleccionado HSDPA, donde el usuario hubiera podido recibir un servicio satisfactorio durante el tiempo que hubiera durado la decisión de selección de RAT. Los resultados obtenidos indican que situaciones como la planteada en el ejemplo anterior son las que provocan un menor nivel de satisfacción de los usuarios cuando las decisiones de selección de RAT se basan en el valor real de CIR/SIR experimentado en el momento de tomar la decisión. De hecho, en todas las ocasiones en las que la decisión tomada por el algoritmo UBReQoS difiere de un caso a otro (un 34 %), en todas ellas la RAT seleccionada a partir del valor de CIR/SIR real es siempre una RAT con menores prestaciones que la seleccionada a partir del valor de CIR/SIR obtenido por el método propuesto, lo cual confirma la afirmación anterior.

Es importante señalar que el método propuesto en este trabajo requiere únicamente almacenar un conjunto de tablas que relacionan el número de canales interferentes activos en cada RAT con un valor de CIR/SIR para cada una de ellas (precalculado según el método propuesto). Estas tablas serían consultadas durante el funcionamiento de la red en lugar de tener que estimar el nivel de interferencia en cada RAT cada vez que sea necesario tomar una decisión de selección de RAT para cada usuario (proceso complejo y computacionalmente costoso). Así pues, frente a la opción consistente en es-

timar el valor real de CIR/SIR experimentado en cada RAT durante el funcionamiento del sistema, el método propuesto en este trabajo no sólo ofrece mayores niveles de satisfacción a los usuarios, según demuestran los resultados obtenidos, sino que también resulta más sencillo y lleva asociado un coste computacional apreciablemente menor.

5.5. Sumario

En el presente capítulo se han mostrado y analizado los resultados obtenidos por simulación para cada uno de los algoritmos propuestos en este trabajo. El rendimiento de dichos algoritmos se ha evaluado de forma comparativa, tanto entre sí como con otros algoritmos sencillos de referencia (alguno de ellos propuesto en la literatura por otros autores). Los algoritmos aportados por el presente trabajo han sido evaluados considerando un amplio conjunto de configuraciones y escenarios de carga con la intención de llevar a cabo un análisis completo del rendimiento de cada uno de ellos. En el capítulo 6 se resumen las principales conclusiones derivadas de dicho análisis.



CAPÍTULO 6

Conclusiones

Mientras la industria de las comunicaciones móviles actualmente dirige sus esfuerzos hacia el desarrollo y explotación comercial de los actuales sistemas de tercera generación, la comunidad investigadora ha comenzado a trabajar durante los últimos años en la definición de lo que serán los futuros sistemas de comunicaciones móviles posteriores a la tercera generación o sistemas de cuarta generación. Algunas iniciativas han sugerido el desarrollo de una nueva tecnología de acceso radio capaz de ofrecer mayores prestaciones binarias que sus predecesoras. No obstante, si bien es cierto que en un futuro se prevé la necesidad de mayores velocidades de transmisión en la interfaz radio, la mayor parte de las actividades de investigación, incluyendo varios proyectos y redes de excelencia recientemente llevados a cabo en el contexto de los Programas Marco de la Unión Europea, parecen coincidir ampliamente en señalar que los futuros sistemas de cuarta generación consistirán en la integración y gestión conjunta de los actuales sistemas celulares, no sólo entre sí, sino también con otras tecnologías radio ya existentes y ampliamente implantadas, así como otras posibles nuevas tecnologías que pudieran surgir en el futuro. Este concepto se basa en la idea de que diferentes redes de acceso radio pueden ser elementos cooperantes dentro de una infraestructura radio heterogénea a través de la cual el proveedor de red puede alcanzar de manera más eficiente los diversos niveles de calidad de servicio requeridos por los usuarios del sistema. Para tal fin, uno de los elementos clave que debe considerar el proveedor de red es la gestión coordinada de los recursos radio (*Common Radio Resource Management*, CRRM) de las diferentes tecnologías de acceso radio como si de un único sistema se tratara, con el objetivo de lograr un aprovechamiento óptimo de los recursos radio disponibles y, en definitiva, un uso más eficiente del sistema en su conjunto.

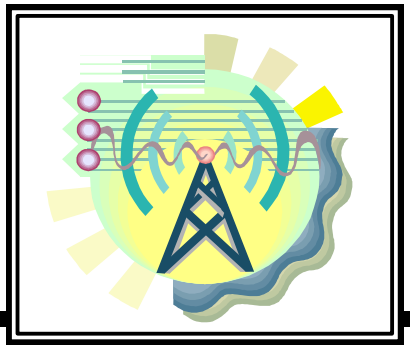
En este contexto, el presente trabajo ha propuesto y evaluado nuevas técnicas CRRM diseñadas para distribuir eficientemente el tráfico de una red radio heterogénea entre las diversas tecnologías de acceso radio (*Radio Access Technologies*, RATs) disponibles, dando respuesta al problema de asignar los usuarios del sistema a aquella red de acceso radio que resulte óptima en base a diversos criterios. El objetivo último de este trabajo es el de proveer soluciones que permitan satisfacer los diversos requisitos de calidad de servicio de los usuarios al mismo tiempo que se garantiza un aprovechamiento óptimo de los recursos radio disponibles en la red.

El primero de los algoritmos propuestos en este trabajo, denominado algoritmo UBET (*Utility Based on Expected Throughput*), pretende ofrecer elevados niveles de calidad de servicio (*Quality of Service*, QoS) a los usuarios del sistema asignándolos a aquella RAT para la cual se prevé la obtención del mejor *throughput*, el cual se estima a partir de los niveles de interferencia experimentados en cada RAT y de la disponibilidad de recursos en cada una de ellas. El rendimiento de este algoritmo se compara, en primer lugar, tomando como referencia el caso en que todos los usuarios de la red son asignados de manera sistemática a aquella RAT con mayores prestaciones binarias. Los resultados obtenidos mostraron que este último planteamiento puede resultar adecuado cuando la carga del sistema es baja o moderada, o cuando las prestaciones ofrecidas por las RATs del sistema difieren entre sí de manera muy considerable, de tal forma que seleccionar otra RAT distinta da como resultado, en última instancia, la obtención de peores niveles globales de QoS. Ante tales situaciones el algoritmo UBET responde asignando la totalidad de usuarios a la RAT con mayores prestaciones binarias. Los resultados mostraron que este modo de proceder permite obtener un elevado nivel de satisfacción global de los usuarios cuando la carga del sistema es baja o moderada. No obstante, si la carga del sistema es elevada, los resultados obtenidos mostraron que los usuarios obtienen mayores niveles de satisfacción cuando son repartidos entre la RAT con mayores prestaciones binarias y otras RATs alternativas, estas últimas operando en modo multicanal, es decir, asignando varios canales de forma simultánea a cada usuario, con la finalidad de ofrecer unas prestaciones binarias comparables a las de la primera. En este caso, la distribución de usuarios realizada según el criterio del algoritmo UBET muestra la obtención de mayores niveles de satisfacción respecto al caso en que los usuarios siguen siendo asignados exclusivamente a la RAT con mayores prestaciones binarias. Ello es debido a que cuando la carga soportada por dicha RAT es lo suficientemente elevada, el nivel de prestaciones ofrecido a los usuarios puede deteriorarse lo suficiente como para que otra RAT alternativa del sistema ofrezca un rendimiento mejor, aun cuando sus prestaciones binarias máximas sean de hecho menores. Ante tal situación el algoritmo UBET reacciona desviando parte de la carga soportada por la RAT más cargada hacia otra RAT alternativa de tal forma que el nivel de satisfacción global de los usuarios experimenta una importante mejora. Así, mientras para cargas bajas o moderadas ambas estrategias obtienen niveles de satisfacción superiores al 90 %, para cargas elevadas el algoritmo UBET obtiene niveles de satisfacción global muy cercanos al 90 % mientras la estrategia de referencia obtiene niveles de satisfacción comprendidos entre el 70 % y el 80 % para los escenarios de carga considerados en este trabajo. Estos resultados ponen de relieve la eficacia del algoritmo UBET en la consecuención del objetivo para el cual fue propuesto.

Una característica del algoritmo UBET es que mientras la carga del sistema no alcanza un determinado nivel, los usuarios son asignados a la RAT con mayores prestaciones binarias, permaneciendo ociosas mientras tanto las demás RATs del sistema. Esta circunstancia implica, por un lado, un bajo grado de utilización de la totalidad de recursos disponibles en el sistema y, por lo tanto, un uso ineficiente de los mismos. Por otra parte, esta manera de proceder puede ocasionar que usuarios de servicios con bajos requisitos de QoS, para los cuales podría resultar suficiente una red con prestaciones modestas, ocupen los recursos de las RATs con mejores prestaciones impidiendo de esta forma que otros usuarios de servicios con elevados requisitos de QoS y que, por lo tanto, necesitan imperativamente las RATs con mejores prestaciones, puedan recibir un servicio satisfactorio.

Para suplir estas carencias del algoritmo UBET, se propone un segundo algoritmo, denominado algoritmo UBReQoS (*Utility Based on Required Quality of Service*). El criterio de este algoritmo consiste en distribuir los usuarios del sistema entre las RATs disponibles en función del nivel de QoS requerido por cada usuario y de las prestaciones ofrecidas por cada RAT del sistema. Frente a otro tipo de soluciones estáticas basadas en asignaciones prefijadas de cada servicio a una determinada RAT, el algoritmo UBReQoS ofrece un elevado grado de flexibilidad adaptando de forma dinámica la distribución del tráfico de la red a las condiciones particulares de demanda experimentadas en cada momento, lo cual resulta en la obtención de mayores niveles de satisfacción para los usuarios, tal y como demuestran los resultados obtenidos. En comparación con el algoritmo UBET se observa que los usuarios de servicios con requisitos de QoS menos exigentes y, por tanto, facturados a un precio menor, experimentan un menor grado de satisfacción a cambio de un significativo incremento en el nivel de satisfacción observado para los usuarios de aquellos servicios más exigentes en términos de QoS y, por lo tanto, facturados a un precio más elevado. Por otra parte, ante un aumento de carga en el sistema, el algoritmo UBReQoS responde de manera más adecuada que el algoritmo UBET, preservando el nivel de satisfacción de los usuarios con mayores requisitos de QoS y que, por tanto, más pagan por sus servicios, a costa de reducir el de los usuarios de servicios menos exigentes, los cuales se facturan a un precio más reducido. En cualquier caso, los resultados obtenidos mostraron la posibilidad de lograr una configuración óptima del algoritmo UBReQoS, mediante un ajuste adecuado de sus parámetros de diseño, gracias a la cual se consigue minimizar el impacto sobre el nivel de satisfacción de aquellos usuarios de servicios menos exigentes. Gracias a la distribución inteligente de usuarios llevada a cabo por el algoritmo UBReQoS, los diversos usuarios del sistema reciben un servicio satisfactorio al mismo tiempo que se realiza un uso eficiente de la totalidad de recursos disponibles en la red.

De cara a una posible implementación práctica de las propuestas realizadas en este trabajo, se ha sugerido un método sencillo para estimar la calidad del enlace radio (dato necesario para la aplicación de los algoritmos propuestos). Frente a la opción consistente en estimar la calidad radio experimentada en cada RAT durante el funcionamiento del sistema, el método propuesto en este trabajo no sólo permite obtener mayores niveles de satisfacción, según demuestran los resultados obtenidos, sino que también resulta más sencillo y lleva asociado un coste computacional considerablemente menor.



APÉNDICE A

Características Relevantes de las Tecnologías de Acceso Radio Consideradas

El interés del presente trabajo se centra, tal y como se ha venido señalando desde el capítulo introductorio, en el estudio, diseño y evaluación de diversos algoritmos para la distribución eficiente del tráfico de una red heterogénea de comunicaciones móviles entre las diferentes alternativas de acceso radio disponibles en dicho entorno heterogéneo. La evaluación de las propuestas realizadas en este trabajo se lleva a cabo mediante el empleo de un sofisticado simulador que, en parte, ha sido desarrollado para los fines de este trabajo. En dicho simulador, descrito en el capítulo 4, se ha considerado una red heterogénea en la cual se integran tres tecnologías de acceso radio diferentes: GPRS (*General Packet Radio Service*), EDGE (*Enhanced Data-rates for GSM/Global Evolution*) y HSDPA (*High Speed Downlink Packet Access*). El presente apéndice pretende proporcionar una breve descripción de aquellos aspectos de mayor relevancia de cada una de estas RATs que han sido considerados e implementados, de forma explícita o implícita, en el simulador a nivel de sistema descrito en el capítulo 4. En este sentido, este apéndice puede considerarse como complementario a dicho capítulo. Los aspectos descritos para cada una de las RATs consideradas son fundamentalmente cuatro. En primer lugar se presenta el método de acceso múltiple empleado por cada tecnología de acceso radio. Luego se describen las características más relevantes para la simulación a nivel de sistema de los diversos modos de transmisión disponibles en cada RAT. A continuación se describe el aspecto temporal de la transmisión de datos a través de la interfaz radio y, por último, el protocolo empleado en cada una de las RATs consideradas para la retransmisión de datos que llegan con error al extremo receptor.

A.1. Características de GPRS

El sistema GPRS [9–11] surgió como una extensión de la red GSM tradicional definida por el ETSI (*European Telecommunications Standardization Institute*) en las especificaciones GSM Fase 2+. El objetivo de GPRS era el de proporcionar nuevos servicios portadores inalámbricos que sirvieran de soporte para la transmisión eficiente de datos. Las principales modificaciones introducidas por GPRS en la interfaz radio fueron, básicamente, la conmutación de paquetes y la aportación de tasas de transmisión más elevadas. Hasta entonces la red GSM, concebida para la transmisión de voz, estaba basada en conmutación de circuitos, lo cual suponía un aprovechamiento ineficiente de los recursos en el caso de transmisión de datos. La introducción de GPRS supone una ruptura respecto a las técnicas anteriores de transferencia de datos en redes de telefonía móvil en la medida en que introduce, de forma novedosa, la conmutación de paquetes para implementar servicios portadores para aplicaciones de datos. Este rasgo hace que GPRS resulte idóneo para dichas aplicaciones, ya que los recursos se asignan sólo cuando se necesita transmitir un paquete de información y se liberan cuando termina la transferencia. De esta forma los recursos se explotan bajo demanda, pudiendo multiplexar varios usuarios sobre un mismo canal físico y mejorando así el aprovechamiento de los recursos radio. Además, el sistema GSM sólo permitía tasas binarias de transmisión de hasta 9.6 kbps, valor bastante modesto para aplicaciones de datos. Con el objetivo de mejorar las capacidades de transmisión de datos en la interfaz radio, el estándar GPRS introdujo diferentes modos de transmisión, cuatro en total, que posibilitaban la transmisión de datos a velocidades superiores a la de GSM.

A.1.1. Acceso múltiple en GPRS

El acceso múltiple define la forma en que el medio radio es accedido, de forma controlada, por los diversos usuarios del sistema. El acceso múltiple de GPRS [194, 195], idéntico al de GSM, consiste en una mezcla de FDMA y TDMA. Además, la técnica de duplexado empleada es FDD, por lo que se utiliza una portadora determinada para la comunicación en sentido ascendente o *uplink* (desde el terminal móvil hacia la estación base) y otra portadora diferente para la comunicación en sentido descendente o *downlink* (desde la estación base hacia el terminal móvil). La figura A.1 muestra la organización del acceso múltiple de GPRS [195].

En la dimensión FDMA/FDD del acceso múltiple se emplean dos bandas separadas con un ancho de 25 MHz cada una de ellas, una para el *uplink* (890 - 915 MHz) y otra para el *downlink* (935 - 960 MHz). Cada banda de frecuencia se divide en 125 canales con un ancho nominal de 200 kHz cada uno. En la práctica se utilizan 124 canales para la transmisión. El canal restante es subdividido en dos mitades de 100 kHz, cada una de las cuales se utiliza como banda de guarda en los límites inferior y superior de ambas bandas de 25 MHz. La función de estas bandas de guarda es dejar un margen de separación espectral para prevenir posibles solapamientos con otras bandas espectrales adyacentes causados por la selectividad limitada de los filtros empleados.

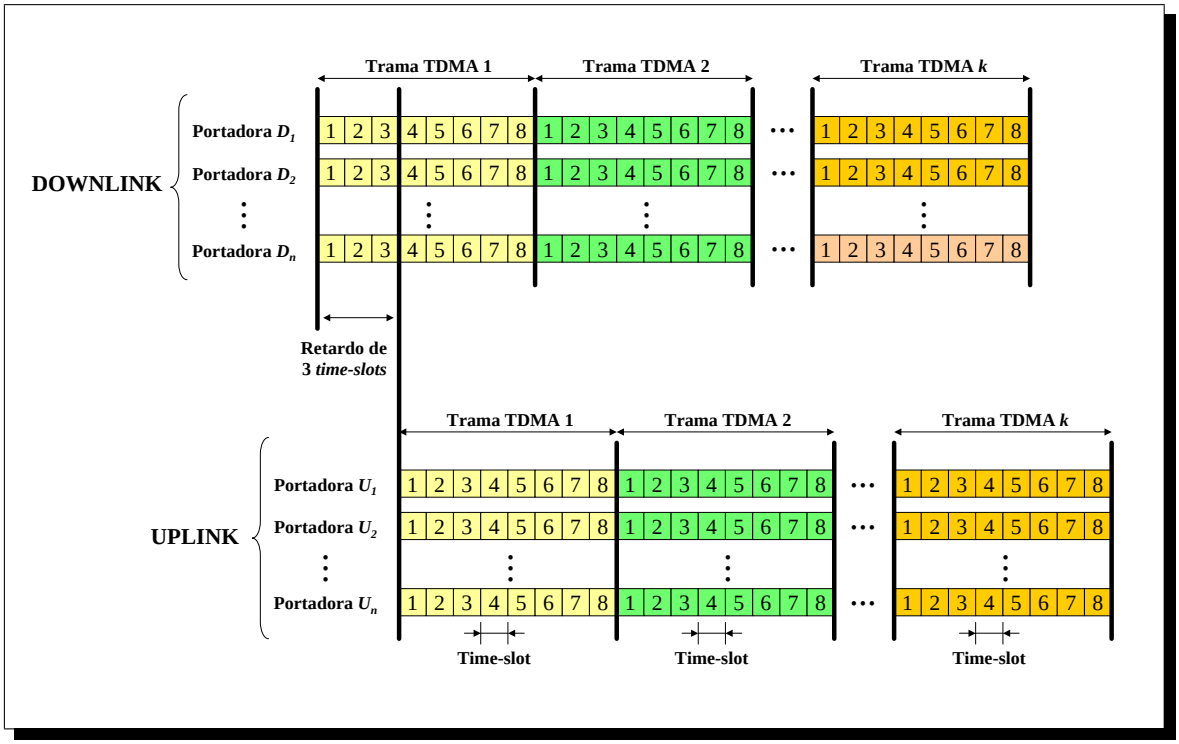


Figura A.1: Acceso múltiple en GPRS.

En la dimensión TDMA del acceso múltiple, cada portadora de radiofrecuencia se divide en ocho intervalos temporales o *time-slots* de aproximadamente 0.577 ms de duración cada uno. Cada *time-slot* se asigna de forma periódica a un usuario determinado. La agrupación de estos ocho *time-slots* constituye la trama TDMA, que se repite de forma periódica y cuya duración es de 4.615 ms (8×0.577 ms). La dimensión temporal incluye otros tipos de agrupaciones temporales más complejas dando lugar a una jerarquía temporal cuya descripción puede encontrarse en [194, 195].

A.1.2. Modos de transmisión en GPRS

Cada uno de los cuatro modos de transmisión disponibles en GPRS se caracteriza por emplear un esquema de codificación de canal (*Coding Scheme*, CS) diferente, aunque todos ellos mantienen la misma modulación GMSK (*Gaussian Minimum Shift Keying*) propia de GSM. Cada uno de ellos ha sido diseñado para ofrecer diferentes grados de compromiso entre robustez frente a errores y tasas de transmisión de información. La tabla A.1 muestra los parámetros de cada modo de transmisión de GPRS. Se puede apreciar que cada modo de transmisión de GPRS aplica un procedimiento de codificación de canal diferente [196] que da lugar a diferentes factores efectivos de tasa de codificación, pero todos ellos mantienen la misma modulación mononivel GMSK [197] que utiliza el sistema GSM. Puesto que el aspecto que varía de un modo de transmisión a otro en GPRS es la codificación de canal, éstos se designan con la nomenclatura CS (*Coding Scheme*). También puede apreciarse que el tamaño del blo-

Modo de transmisión	Tipo de modulación	Factor del código	Bits de datos por bloque radio	Tasa (kbps)
CS-1	GMSK	1/2	181	9.05
CS-2	GMSK	$\approx 2/3$	268	13.4
CS-3	GMSK	$\approx 3/4$	312	15.6
CS-4	GMSK	1	428	21.4

Tabla A.1: Modos de transmisión de GPRS [198].

que de datos tomado de las capas superiores varía en función del modo de transmisión que vaya a ser utilizado en la transmisión radio, desde los 181 bits para CS-1 hasta los 428 bits para CS-4. En cualquier caso, independientemente del modo de transmisión empleado, a la salida del codificador de canal siempre se obtienen bloques codificados (bloques radio) con una longitud constante de 456 bits.

El aspecto de mayor relevancia a señalar es que reduciendo la cantidad de bits de datos en cada bloque radio (de tamaño constante), puede aumentarse la cantidad de bits redundantes utilizados para la detección y corrección de los errores de transmisión radio, aumentando así la robustez ante errores. Este hecho se ilustra en la figura A.2, obtenida como resultado de varias simulaciones a nivel de enlace. En dicha figura se muestra el valor del BLER (parámetro que representa la proporción de bloques RLC erróneos, ver sección A.1.3) en función de la relación portadora a interferencia (*Carrier-to-Interference Ratio*, CIR). Según se aprecia en dicha figura, para un mismo valor de relación CIR, CS-1 es el modo de transmisión con el cual se obtiene una menor proporción de información errónea, mientras que CS-4 es el menos robusto.

Sin embargo, al reducir la cantidad de bits de datos en cada bloque radio, la velocidad de transmisión de datos, mostrada en la última columna de la tabla A.1, es también menor, lo cual se muestra en la figura A.3. En dicha figura se representa el *throughput*, definido como la cantidad de bits correctamente transmitidos por unidad de tiempo, en función de la relación CIR. Las curvas mostradas en la figura A.3 se obtienen evaluando, para cada modo de transmisión, la expresión

$$Throughput = R \cdot (1 - BLER) \tag{A.1}$$

donde R representa la velocidad binaria de transmisión y corresponde al máximo valor de *throughput* que puede obtenerse (cuando todos los bits transmitidos se reciben sin error). El valor de R se obtiene dividiendo la cantidad de bits de datos transmitidos en un bloque radio entre la duración de dicho bloque, que en el caso de GPRS es de 20 ms según se verá en la sección A.1.3. Puede apreciarse, así pues, que cada modo de transmisión de GPRS ofrece diferente grado de compromiso entre protección ante errores y velocidad de transmisión de datos. Esta es una característica común a cualquier interfaz radio en la cual se disponen de diferentes modos de transmisión y que permite adaptar el proceso de transmisión a la calidad experimentada en el enlace radio.

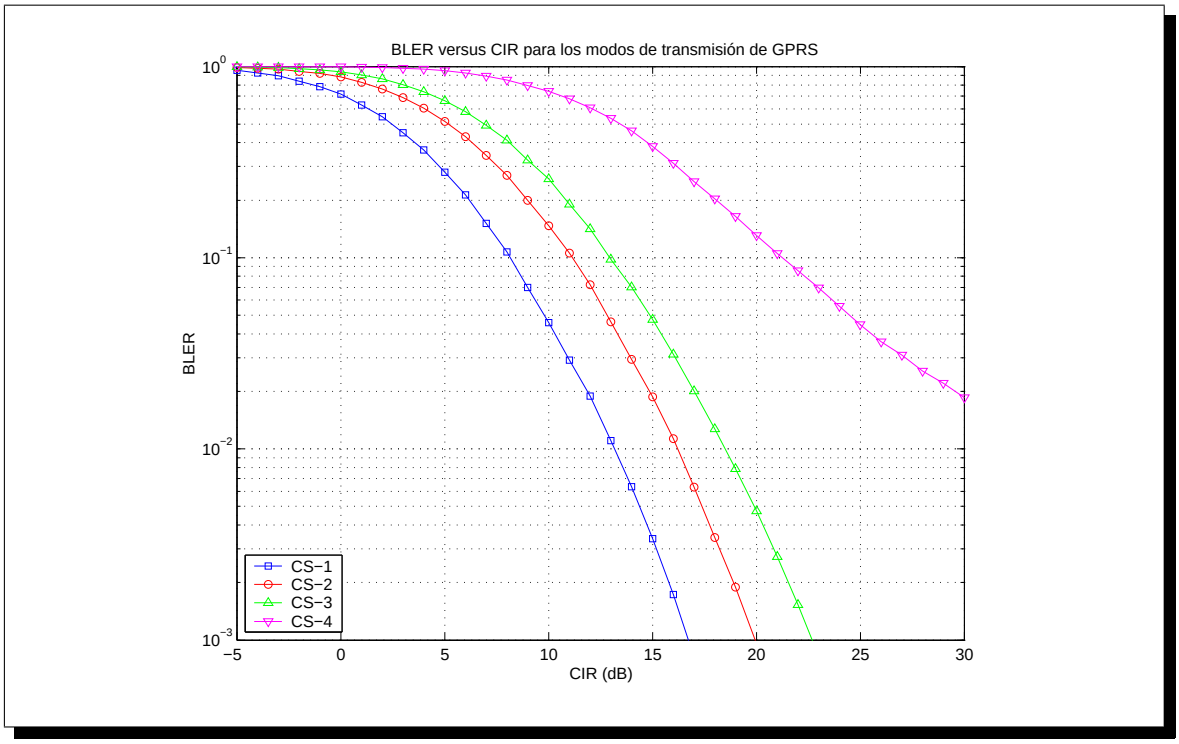


Figura A.2: BLER versus CIR para los modos de transmisión de GPRS [187].

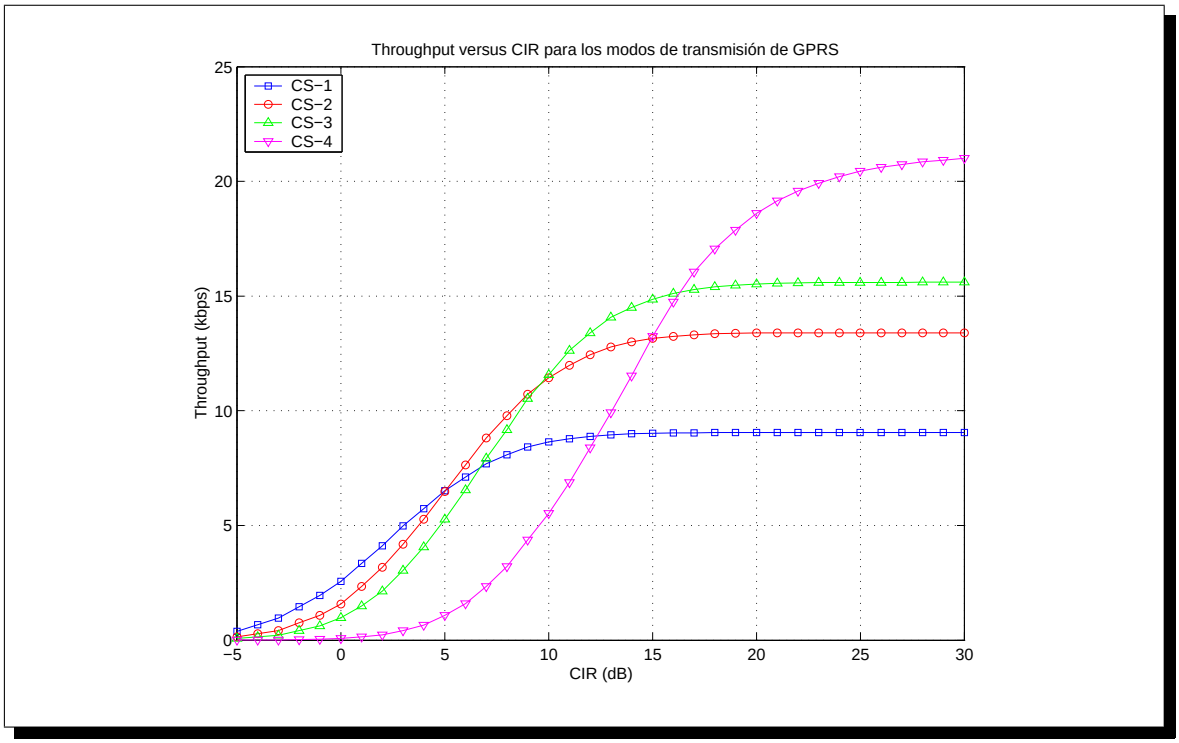


Figura A.3: *Throughput* versus CIR para los modos de transmisión de GPRS.

A.1.3. Transmisión de datos en GPRS

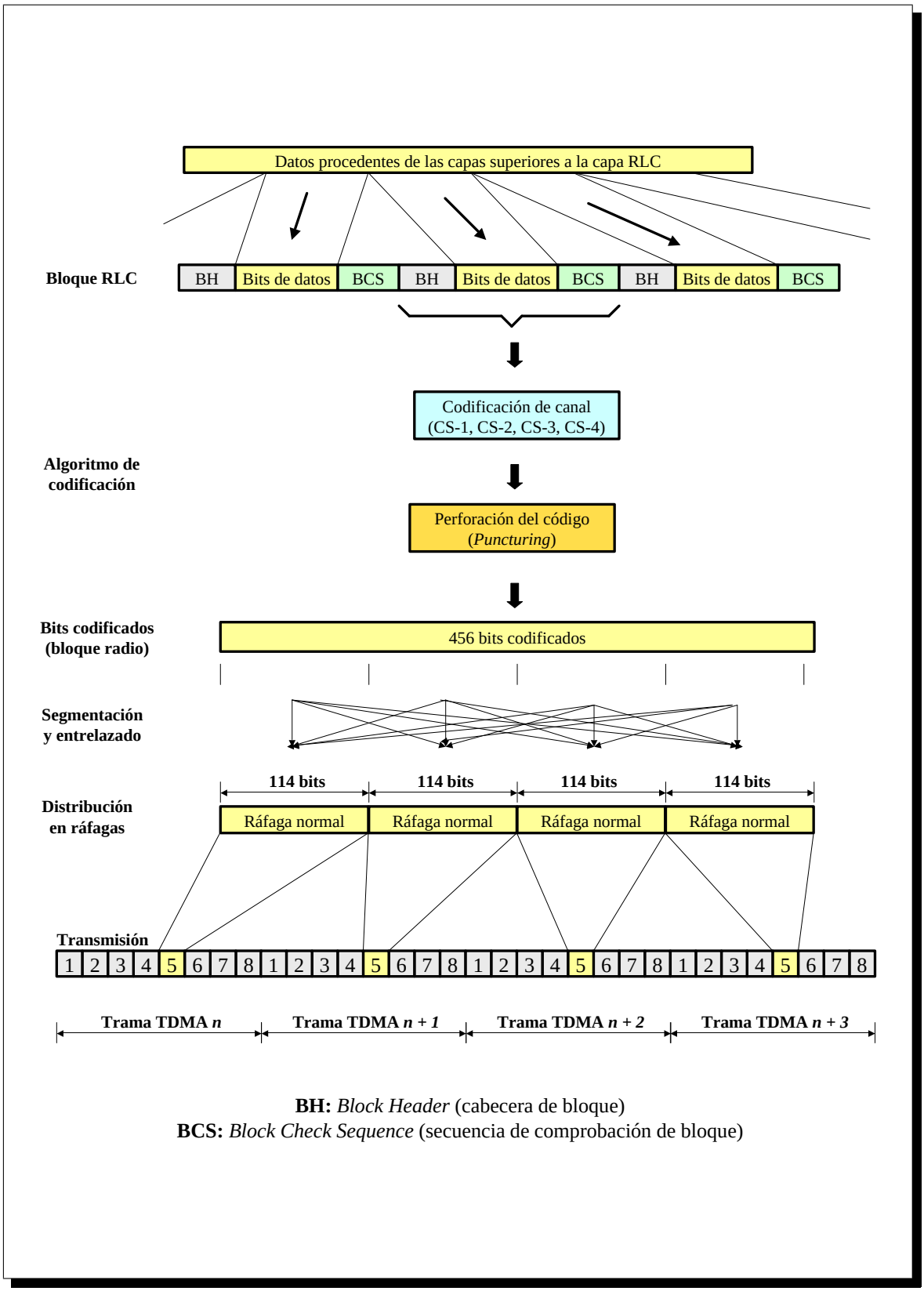
Cuando la información de usuario procedente de las capas altas de la torre de protocolos llega a la capa RLC (*Radio Link Control*) de GPRS [199], capa perteneciente al nivel de enlace de datos según el modelo de referencia OSI (*Open System Interconnection*), se segmenta en unidades que reciben el nombre de *bloques RLC* [200]. Cada bloque RLC posee un campo de datos cuyo tamaño depende del modo de transmisión a emplear (ver tabla A.1) así como cierta información adicional propia de la capa RLC que se incluye en los campos de cabecera (*Block Header*, BH) y de comprobación de errores (*Block Check Sequence*, BCS). En el propio proceso de codificación de canal se aplican ciertos patrones de *puncturing* o "perforación" que eliminan ciertos bits para ajustar la longitud de la secuencia resultante a un valor constante de 456 bits.

Sobre estos 456 bits se aplica un proceso de segmentación y entrelazado consistente en reordenar los bits dentro del bloque radio codificado. La razón por la cual se realiza esta operación es para mejorar la capacidad de corrección de errores. Las técnicas de corrección soportan un número bastante limitado de bits erróneos dentro de una determinada secuencia y resultan menos eficaces a medida que éstos están más próximos entre sí. Desafortunadamente, en el ámbito práctico, los errores no se producen en bits aislados, sino que suelen producirse en agrupaciones de bits consecutivos, llegando a hacer inútiles las técnicas de corrección. El entrelazado representa una mejora a este problema. Los bits, antes de ser transmitidos, son distribuidos en diferentes ráfagas de forma salteada (la descripción detallada del proceso puede encontrarse en [196]). De esta forma, los bits erróneos consecutivos producidos por la transmisión radio se dispersan al reordenar los bits en el receptor para recuperar la secuencia original. El resultado es que, en la secuencia reordenada, los bits erróneos quedan distribuidos con cierta separación entre sí. Sobre esta secuencia es posible aplicar de forma más eficaz las técnicas de corrección de errores.

Los 456 bits resultantes después del proceso de entrelazado se dividen en cuatro segmentos de 114 bits, cada uno de los cuales se acomoda en una ráfaga normal (*normal burst*). Cada uno de estos cuatro segmentos de 114 bits se reparten a lo largo de cuatro tramas TDMA consecutivas, tal y como muestra la figura A.4, acomodando los datos en el mismo *time-slot* relativo que tiene asignado el usuario en cuestión. Así pues, la transmisión de cada bloque radio se realiza en cuatro tramas TDMA consecutivas.

A.1.4. Protocolo de retransmisión en GPRS

El estándar GPRS define un mecanismo de retransmisión de información recibida con error a nivel de bloque RLC. Este mecanismo se basa en un procedimiento ARQ (*Automatic Repeat reQuest*) de ventana deslizante con rechazo selectivo mediante el cual el extremo receptor solicita al transmisor la retransmisión de aquellos bloques RLC que han sido recibidos con error. La especificación completa y detallada se puede encontrar en [200]. La figura A.5 ilustra el principio de funcionamiento.



BH: *Block Header* (cabecera de bloque)
BCS: *Block Check Sequence* (secuencia de comprobación de bloque)

Figura A.4: Transmisión de datos en GPRS.

El método de retransmisión se apoya en la numeración sucesiva de los bloques RLC transmitidos a través de la interfaz radio. Cada bloque RLC tiene asignado un número de secuencia (*Block Sequence Number*, BSN) cuyo valor es una unidad superior al del bloque RLC precedente. El valor del BSN se indica mediante un campo de 7 bits, lo que significa que es posible utilizar hasta un total de $2^7 = 128$ números distintos, dando lugar a un espacio de números de secuencia (*Sequence Number Space*, SNS) con un tamaño de 128 elementos (SNS = 128). El valor de los BSNs varía desde 0 hasta 127 y cada 128 bloques RLC éstos vuelven a numerarse empezando desde cero, tal y como se muestra en la figura A.5.

El protocolo ARQ empleado se caracteriza por el empleo de ventanas, tanto en el extremo transmisor como en el extremo receptor. En el extremo transmisor, la ventana de transmisión se define como el conjunto de bloques RLC cuyo BSN cumple la condición [V(A) $\leq$ BSN $<$ V(S)] mod SNS. Dicha ventana posee un tamaño variable delimitado por dos variables de estado V(A) y V(S) que cumplen que [V(S) - V(A)] mod SNS $\leq$ WS, siendo WS el tamaño de ventana (*Window Size*), que en GPRS es igual a 64. La variable V(A) contiene el valor BSN del bloque RLC más antiguo del cual todavía no se ha recibido una confirmación positiva por parte del receptor, es decir, que el receptor todavía no ha confirmado como bloque recibido correctamente (sin error). La variable V(A) puede tomar valores en el intervalo [0, SNS-1]. El valor de V(A) se establece a cero al inicio de cada nueva transmisión y su valor se actualiza cada vez que se recibe un mensaje del extremo receptor indicando aquellos bloques RLC recibidos con y sin error. La variable de estado V(S) contiene el valor BSN del siguiente bloque RLC a transmitir, pudiendo tomar valores en el intervalo [0, SNS-1]. El valor de V(S) se establece a cero al inicio de cada nueva transmisión y su valor se incrementa en una unidad cada vez que se transmite un nuevo bloque RLC. Es decir, al transmitir el bloque con BSN = V(S) entonces V(S) se incrementa en una unidad. El valor de V(S) no puede exceder en ningún momento el valor de V(A) mod SNS en una cantidad superior a WS, es decir, [V(S) - V(A)] mod SNS $\leq$ WS.

Durante la transmisión, el transmisor debe transmitir siempre en primer lugar el bloque RLC más antiguo que haya sido confirmado por el receptor como recibido con error, luego el siguiente y así sucesivamente. Cuando no hay bloques RLC confirmados negativamente entonces pueden darse dos casos. El primero es que [V(S) $<$ V(A) + WS] mod SNS, es decir, la ventana no ha alcanzado el tamaño máximo, en cuyo caso se procede a la transmisión de un nuevo bloque RLC, aquel con BSN = V(S). Cuando no quedan bloques RLC por transmitir, se retransmiten aquellos bloques RLC pendientes de confirmación. Así el transmisor no está ocioso y si alguno de los bloques pendientes de confirmación fue inicialmente recibido con error, estas retransmisiones hacen llegar nuevamente los bloques RLC antes de que el receptor solicite su retransmisión, pudiendo agilizar el proceso. Estas retransmisiones se repiten hasta haber recibido un mensaje de confirmación por parte del receptor. El segundo caso posible cuando no quedan bloques RLC confirmados negativamente es que V(S) = [V(A) + WS] mod SNS, es decir, la ventana del transmisor alcanza su tamaño máximo permitido, en cuyo caso el transmisor procede también a retransmitir los bloques RLC pendientes de confirmación mientras se cumpla la condición V(S) = [V(A) + WS] mod SNS.

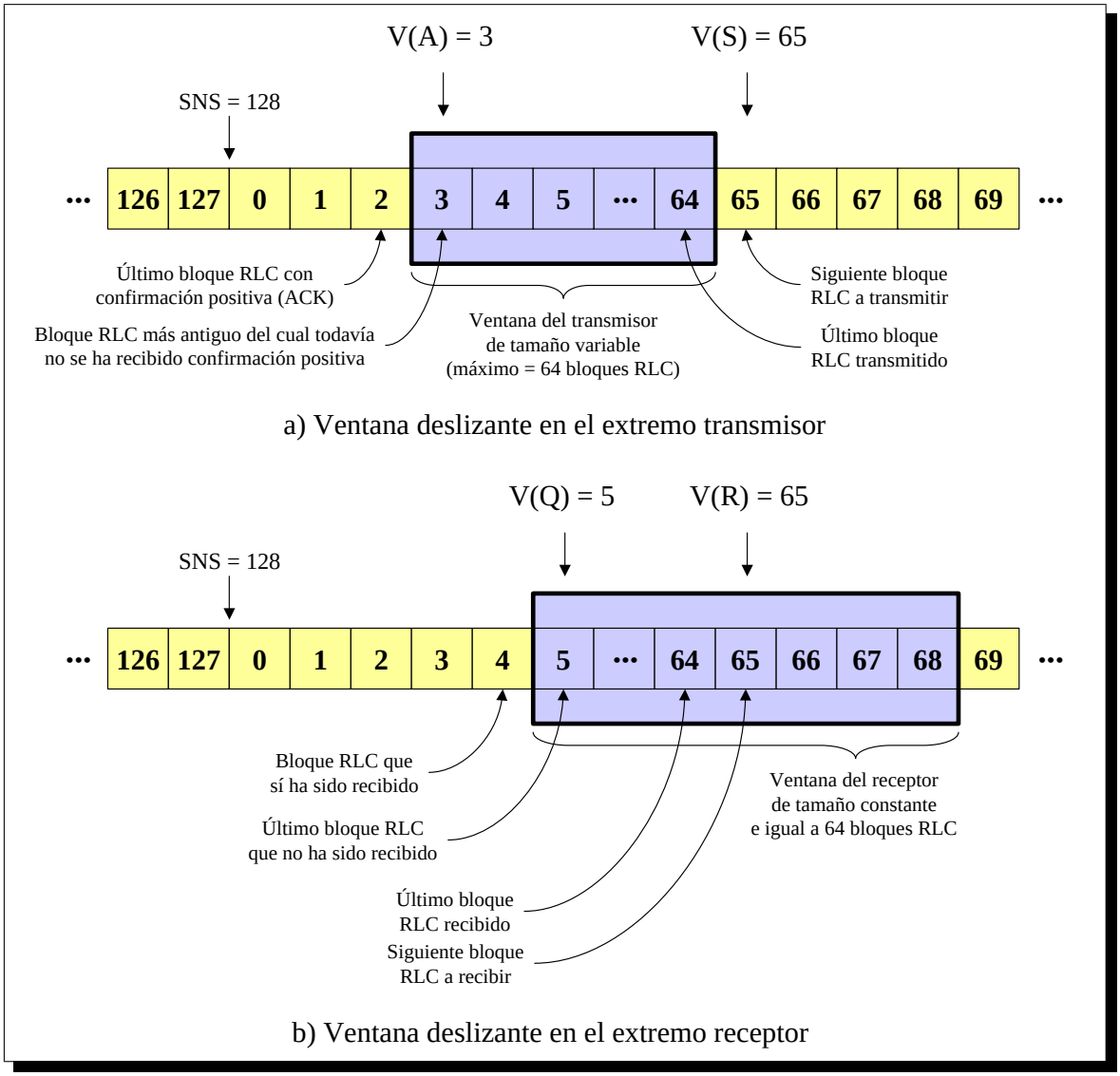


Figura A.5: Retransmisión de datos en GPRS.

En el extremo receptor, la ventana de recepción se define como el conjunto de bloques RLC cuyo BSN cumple la condición [V(Q) ≤ BSN < V(Q) + WS] mod SNS. Dicha ventana posee un tamaño constante de WS bloques RLC (64 en GPRS) y viene determinada por el valor de la variable de estado V(Q). La variable V(Q) contiene el BSN más bajo todavía no recibido. La variable V(Q) puede tomar cualquier valor en el intervalo [0, SNS-1]. Se establece a cero al comienzo de cada nueva transmisión. Su valor se actualiza al recibir el bloque RLC cuyo BSN es igual a V(Q), tomando entonces el valor BSN del siguiente bloque RLC en la ventana del receptor que todavía no ha sido recibido o tomando el valor de la variable de estado V(R) si todos los bloques RLC de la ventana receptora han sido recibidos. La variable de estado V(R) indica el BSN cuyo valor es una unidad mayor que el máximo valor de BSN recibido, pudiendo tomar valores en el intervalo [0, SNS-1]. Su valor se establece a cero al inicio de una nueva transmisión y se actualiza con cada nuevo bloque RLC recibido.

A.2. Características de EDGE

El sistema GPRS supuso un avance en la transmisión de datos respecto a las tecnologías celulares predecesoras, introduciendo la conmutación de paquetes para un aprovechamiento más eficiente de los recursos radio e incrementando las tasas binarias de transmisión en la interfaz radio mediante la introducción de diversos modos de transmisión. No obstante, en el proceso evolutivo hacia los sistemas 3G se estableció la necesidad de velocidades de transmisión de 144 kbps (vehículos) y de 384 kbps (peatones) en entornos exteriores (*outdoor*) y de hasta 2 Mbps en entornos interiores (*indoor*). Las tasas binarias ofrecidas por GPRS resultaban claramente insuficientes. La máxima velocidad teórica que puede alcanzar un usuario bajo GPRS corresponde a una situación ideal en la cual la calidad del enlace radio permite utilizar CS-4 sin experimentar errores de transmisión (21.4 kbps por *time-slot*) y dicho usuario tiene asignados los 8 *time-slots* de la trama TDMA, resultando en una velocidad de $8 \times 21.4 = 171.2$ kbps. El sistema EDGE (*Enhanced Data-rates for GSM/Global Evolution*) [12–14] surgió como el siguiente paso lógico en el proceso evolutivo hacia los sistemas 3G, posibilitando velocidades de transmisión radio superiores a las disponibles hasta entonces mediante la introducción de nuevos modos de transmisión.

A.2.1. Acceso múltiple en EDGE

El acceso múltiple en EDGE se basa en una mezcla de FDMA/FDD y TDMA, al igual que en GSM y GPRS. De hecho, las bandas de frecuencias empleadas y sus subdivisiones en canales de radiofrecuencia así como las estructuras de tramas y otras agrupaciones de mayor y menor nivel jerárquico son idénticas a las de GPRS [194, 195].

A.2.2. Modos de transmisión en EDGE

Con EDGE aparecen nuevos modos de transmisión capaces de ofrecer velocidades de transmisión más elevadas. De forma similar a GPRS, cada modo de transmisión se basa en un procedimiento de codificación de canal diferente [196], aunque como novedad, los nuevos modos de EDGE ya no sólo emplean la modulación mononivel GMSK de GSM y GPRS, sino que ahora también utilizan una modulación multinivel 8PSK (8 *Phase Shift Keying*) [197], gracias a la cual pueden alcanzar velocidades de transmisión más elevadas. La tabla A.2 resume los parámetros más relevantes de cada modo de transmisión y las figuras A.7 y A.8 muestran sus prestaciones. Una característica interesante de los modos de transmisión de EDGE es que éstos se agrupan en cuatro familias cada una de las cuales posee una unidad básica de datos con diferente tamaño, según muestra la figura A.6. Dentro de una misma familia se obtienen diferentes factores de código transmitiendo un número diferente de unidades de datos dentro de un bloque radio. Estas familias se crean para posibilitar la retransmisión de información errónea con modos de transmisión más robustos que el inicial (lo cual no es posible en GPRS), utilizando modos de la misma familia pero con menos unidades de datos.

Modo de transmisión	Tipo de modulación	Factor del código	Tipo de familia	Bits de datos por bloque radio	Tasa (kbps)
MCS-1	GMSK	0.53	C	$1 \times 176 = 176$	8.8
MCS-2	GMSK	0.66	B	$1 \times 224 = 224$	11.2
MCS-3	GMSK	0.85	A *padding*	$1 \times 272 = 272$	13.6
			A	$1 \times 296 = 296$	14.8
MCS-4	GMSK	1.00	C	$2 \times 176 = 352$	17.6
MCS-5	8PSK	0.37	B	$2 \times 224 = 448$	22.4
MCS-6	8PSK	0.49	A *padding*	$2 \times 272 = 544$	27.2
			A	$2 \times 296 = 592$	29.6
MCS-7	8PSK	0.76	B	$4 \times 224 = 896$	44.8
MCS-8	8PSK	0.92	A *padding*	$4 \times 272 = 1088$	54.4
MCS-9	8PSK	1.00	A	$4 \times 296 = 1184$	59.2

Tabla A.2: Modos de transmisión de EDGE [198].

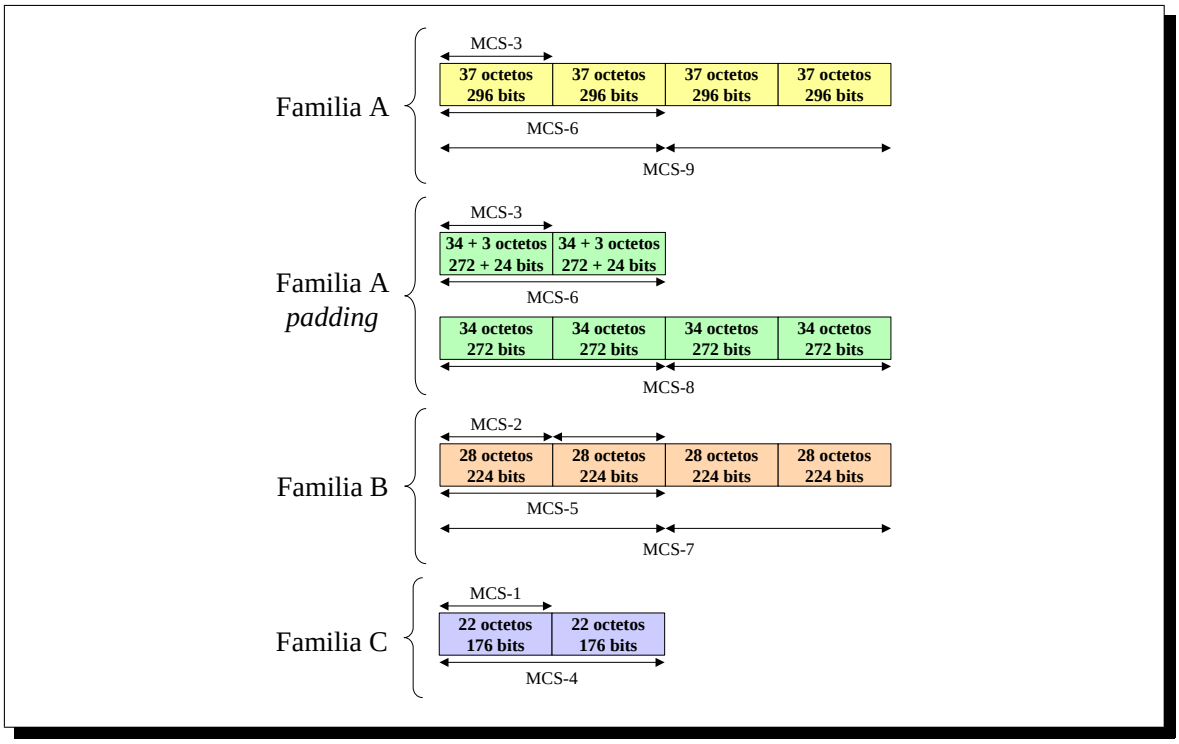


Figura A.6: Familias de modos de transmisión de EDGE.

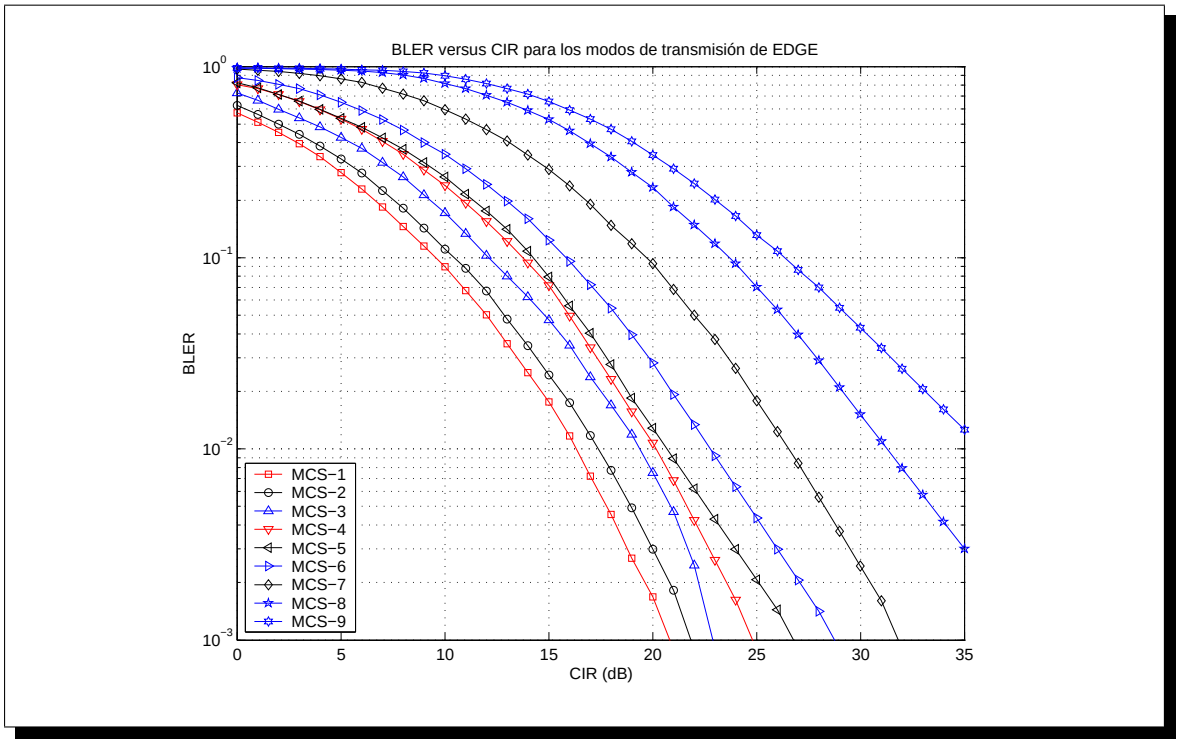


Figura A.7: BLER versus CIR para los modos de transmisión de EDGE [3].

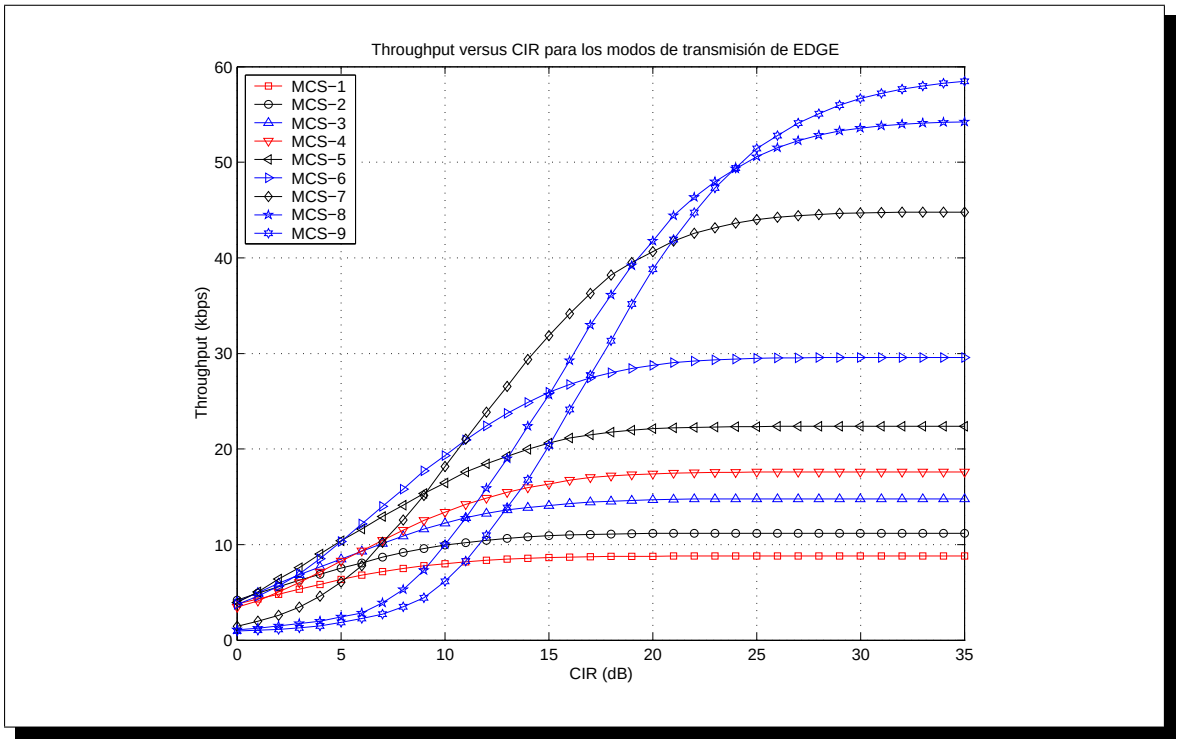


Figura A.8: *Throughput* versus CIR para los modos de transmisión de EDGE.

A.2.3. Transmisión de datos en EDGE

El proceso de transmisión de datos de EDGE es bastante similar al mostrado en la figura A.4 para el caso de GPRS, aunque con ciertas particularidades. En primer lugar, la codificación de canal realizada y el patrón de *puncturing* empleado dependen de cuál de los modos de transmisión de la tabla A.2 vaya a ser utilizado en la transmisión. El resultado de dicho proceso no es siempre una secuencia de 456 bits como sucedía con GPRS. Ahora esta longitud puede ser de 464 bits si el modo de transmisión se basa en la modulación mononivel GMSK (1 bit por símbolo modulado) o de $464 \times 3 = 1392$ bits si el modo de transmisión se basa en la modulación multinivel 8PSK (3 bits por símbolo modulado). En ambos casos se aplica un determinado esquema de entrelazado descrito en [196] y los bits se dividen en cuatro segmentos (de 116 bits para GMSK y de 348 bits para 8PSK) cada uno de los cuales se acomoda en una ráfaga normal (la longitud y el formato de la ráfaga normal para GMSK y 8PSK es diferente, ver [195]).

Al igual que en GPRS, cada bloque radio se transmite en cuatro ráfagas normales sobre cuatro tramas TDMA consecutivas, empleando para ello un total de 20 ms. Sin embargo, a diferencia de GPRS, en EDGE un mismo bloque radio puede contener dos bloques RLC diferentes. Esto sucede cuando se transmiten cuatro unidades de datos, es decir, cuando se emplea MCS-7, MCS-8 o MCS-9, en cuyo caso las cuatro unidades de datos se dividen a nivel lógico en dos bloques RLC diferentes, cada uno con un BSN diferente, etc. Con MCS-7 estos dos bloques RLC son entrelazados sobre las cuatro ráfagas que integran un bloque radio, de manera que si cualquiera de dichas ráfagas contiene bits erróneos, entonces los dos bloques RLC se reciben con error. Con MCS-8 y MCS-9, en cambio, los dos bloques RLC transmitidos en un bloque radio son entrelazados sobre dos ráfagas diferentes, de manera que si una de las ráfagas sufre un error, dicho error sólo afectará a uno de los bloques RLC pero el otro todavía podrá ser decodificado correctamente. Esto se hace así porque MCS-8 y MCS-9 tienen factores de código prácticamente igual a 1 en ambos casos, lo cual significa que son poco robustos ante los errores de transmisión, y además se basan en una modulación multinivel que resulta más susceptible de sufrir errores que una modulación mononivel. En el resto de modos de transmisión, desde MCS-1 hasta MCS-6, cada bloque radio contiene un único bloque RLC que es entrelazado sobre las cuatro ráfagas normales del bloque radio.

A.2.4. Protocolo de retransmisión en EDGE

El estándar EDGE define un mecanismo de retransmisión de información recibida con error a nivel de bloque RLC. Este mecanismo se basa en el mismo procedimiento ARQ de ventana deslizante con rechazo selectivo en el cual se basa GPRS [200]. No obstante, existen ciertas diferencias. Por un lado, el valor de BSN de cada bloque RLC se indica empleando para ello 11 bits, lo cual da lugar a un espacio de números de secuencia con un tamaño igual a $2^{11} = 2048$ elementos, es decir, SNS = 2048. Por otro lado, el tamaño máximo de ventana WS no es fijo, sino que depende de la cantidad de *time-slots* dentro de una misma trama TDMA que son asignados de manera simultánea a un mismo usuario. Cuando se asigna un *time-slot* por usuario, entonces WS = 192.

A.3. Características de HSDPA

Actualmente estamos asistiendo al despliegue de la tecnología UMTS con la cual es posible alcanzar elevadas velocidades de transmisión, teóricamente de hasta 2 Mbps. Ello posibilita a los proveedores la prestación de diversos servicios multimedia de banda ancha como, por ejemplo, la videoconferencia, que hasta hace no mucho tiempo eran tecnológicamente inviables en un sistema celular. A pesar de ello, los proveedores apuestan por servicios cada vez más complejos y sofisticados que requieren cada vez mayores velocidades de transmisión y menores retardos. Para posibilitar este tipo de servicios de altas prestaciones surge en la *Release 5* de las especificaciones del 3GPP la tecnología HSDPA (*High Speed Downlink Packet Access*) [201–205]. Suele decirse que el sistema HSDPA, en ocasiones referido como sistema 3.5G, es a UMTS lo que GPRS es a GSM. A grandes rasgos, HSDPA es una tecnología pensada para ofrecer elevadas tasas binarias y bajos retardos de transmisión en el enlace descendente (desde la estación base hacia el terminal móvil), basada en el uso de un canal descendente compartido de alta velocidad denominado HS-DSCH (*High Speed Downlink Shared CHannel*) que mejora ciertos aspectos de su homólogo DSCH (*Downlink Shared CHannel*) existente en UMTS *Release 99*. El estándar HSDPA introduce nuevos modos de transmisión que permiten velocidades máximas teóricas de hasta 10 Mbps pero, además, introduce otras nuevas técnicas respecto a UMTS pensadas para lograr un aprovechamiento más eficiente de los recursos radio [206, 207]. En HSDPA *Release 5* aparecen las técnicas conocidas como *Adaptive Modulation and Coding* (AMC), *Hybrid Automatic Repeat reQuest* (HARQ) y *fast scheduling* y en versiones posteriores se introducen las técnicas *Multiple Input Multiple Output* (MIMO) y *Fast Cell Selection* (FCS). En las secciones siguientes se describirán únicamente las técnicas incluidas en HSDPA *Release 5*.

A.3.1. Acceso múltiple en HSDPA

El acceso múltiple en HSDPA guarda cierta relación con el definido por la tecnología WCDMA (*Wideband Code Division Multiple Access*) [166] en la cual se basa UMTS, donde se define un acceso múltiple basado en CDMA de secuencia directa (*Direct-Sequence Code Division Multiple Access*, DS-CDMA). Los bits de información se multiplican por una secuencia aleatoria de bits (denominados *chips*) con una tasa mucho mayor que la señal de datos, dando lugar a un ensanchamiento espectral de la señal original. La tasa de esta secuencia aleatoria, denominada *código de ensanchamiento* o *spreading code* depende del modo de operación o duplexado empleado.

El sistema HSDPA, al igual que UMTS, puede operar en modo FDD [208–212] o en modo TDD [213–217]. La componente FDD hace uso de dos bandas de frecuencia emparejadas de 60 MHz cada una, una para el *uplink* (1920 - 1980 MHz) y otra para el *downlink* (2110 - 2170 MHz). La componente TDD utiliza dos bandas de frecuencia no emparejadas, una de 15 MHz que va desde los 2010 MHz hasta los 2025 MHz y otra de 20 MHz que va desde los 1900 MHz hasta los 1920 MHz.

La componente FDD opera con una tasa de chip de 3.84 Mcps, lo cual da lugar a un ancho de banda de aproximadamente 5 MHz. Esto se traduce en la existencia de 12 portadoras FDD (estas portadoras no se utilizan para diferenciar las transmisiones de varios usuarios mediante FDMA). La componente TDD puede operar con una tasa de chip también de 3.84 Mcps (dando lugar a 7 portadoras TDD) o con una tasa de chip de 1.28 Mcps. En este último caso el ancho de banda de transmisión es de aproximadamente 1.6 MHz. Tanto en FDD como en TDD, el acceso múltiple responde, a grandes rasgos, al esquema mostrado en la figura A.9. En dicho esquema se puede distinguir entre aspectos relativos al dominio del tiempo y al dominio del código [218].

En el dominio del tiempo se define una trama con una duración de 10 ms la cual se divide en 15 *slots* de 667 μs cada uno. Esta división es característica de UMTS, aunque en HSDPA también se define para el canal compartido HS-DSCH una subtrama con una duración de 2 ms, la cual se divide a su vez en 3 *slots* también de 667 μs cada uno. En FDD, la división de una trama o una subtrama en *slots* no se utiliza para separar usuarios tal y como se hace en TDMA, sino que simplemente sirve para determinar la frecuencia con la cual se actualiza el valor de ciertas variables y parámetros de transmisión en la interfaz radio. En cambio, en TDD los *slots* sí que se utilizan para separar usuarios dentro de una misma trama de 10 ms. Por lo tanto, en TDD sí que existe cierta componente TDMA en el acceso múltiple, además de la componente DS-CDMA. A partir de estas divisiones temporales se define el intervalo de tiempo de transmisión (*Transmission Time Interval*, TTI) que determina la periodicidad con la cual la capa física recibe bloques de datos de la capa superior. En UMTS se permiten TTIs de 10 ms, 20 ms, 40 ms y 80 ms. En HSDPA FDD el valor del TTI es fijo y se ha reducido a 2 ms (duración de una subtrama) para disminuir el tiempo de espera en la retransmisión de datos. Para la componente HSDPA TDD se contemplan valores de 10 ms (3.84 Mcps TDD) y de 5 ms (1.28 Mcps TDD).

En el dominio del código cabe señalar la distinción entre dos tipos de códigos [210, 215]. En primer lugar la señal en banda base se multiplica por una secuencia o código de canalización (*channelisation code*) cuya tasa se corresponde con la tasa de chip (1.28 ó 3.84 Mcps), utilizada para ensanchar el ancho de banda de la señal inicial. Estos códigos de canalización son ortogonales entre sí para los usuarios de una misma celda, evitando (teóricamente) la interferencia entre dichos usuarios. Las señales de los usuarios de una misma celda se suman y seguidamente se efectúa una segunda multiplicación de la señal resultante por una secuencia o código de aleatorización (*scrambling code*), diferente para cada estación base. La secuencia de aleatorización también posee la tasa de chip correspondiente pero no produce ensanchamiento alguno en el espectro. Su misión es la de asegurar que las transmisiones de dos estaciones base diferentes se comportan como señales aleatorias, separando así las transmisiones de células diferentes (aunque con cierta interferencia entre estaciones bases cercanas ya que estos códigos pseudo-aleatorios no son perfectamente ortogonales entre sí). Cada célula dispone de varios códigos de aleatorización. Cuando se agotan los códigos de canalización disponibles, éstos son reutilizados dentro de la misma celda pero empleando entonces un código de aleatorización distinto para cada uno de ellos. En este caso la interferencia teórica entre usuarios de una misma celda no es completamente nula.

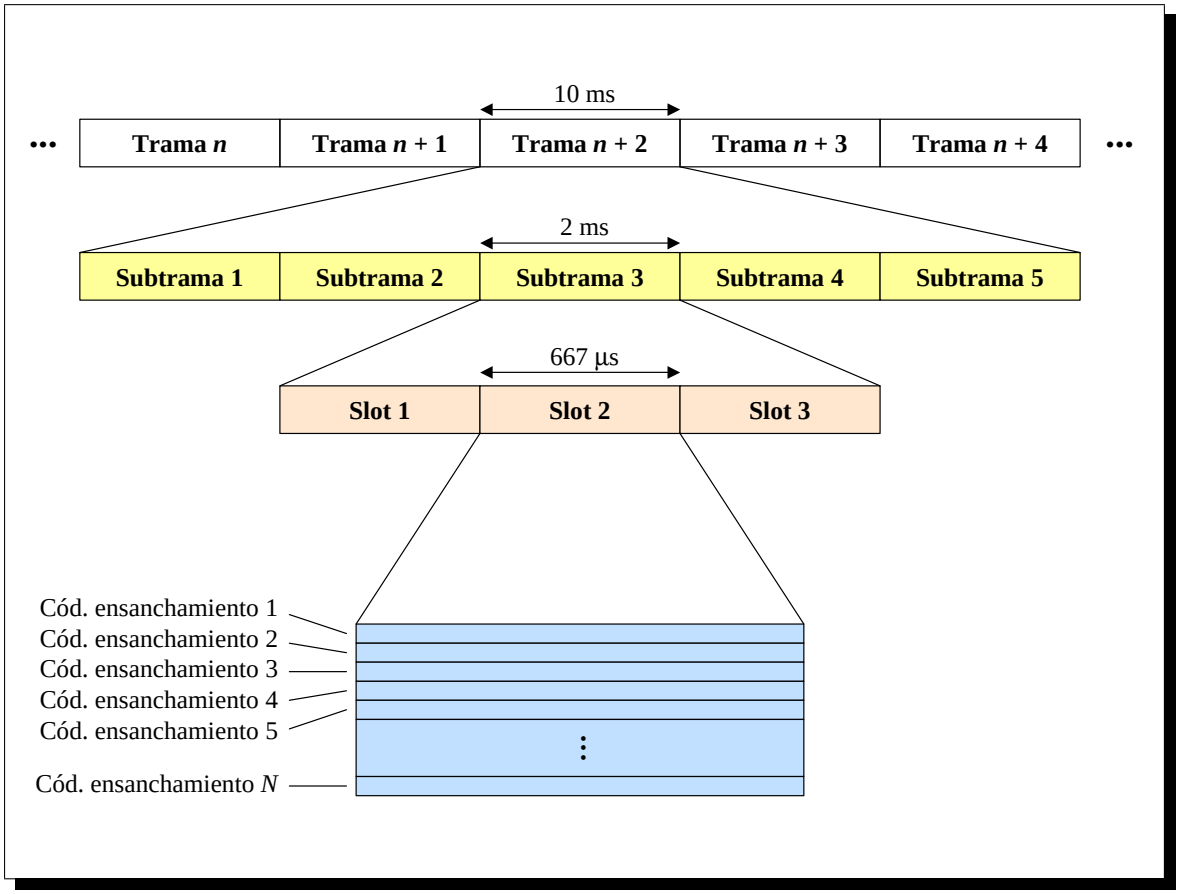


Figura A.9: Acceso múltiple en HSDPA.

La transmisión HS-DSCH en FDD utiliza códigos de canalización con un factor de ensanchamiento (*spreading factor*) constante e igual a 16. Esto permite la existencia de hasta 16 códigos de canalización distintos por celda, aunque por lo general sólo una parte de ellos se reserva para su utilización en la transmisión a través del canal HS-DSCH de HSDPA. La multiplexación de varios terminales móviles dentro de un mismo TTI se realiza en el dominio del código, asignando a cada uno de ellos un código de canalización diferente. También se permite la transmisión multi-código, lo que significa que un mismo terminal móvil de usuario puede tener asignados varios códigos de canalización de manera simultánea en el mismo TTI. En función de la capacidad de dicho terminal, el número máximo de códigos simultáneos puede ser 5, 10 ó 15 (todos ellos deben emplear el mismo código de aleatorización). La transmisión HS-DSCH en TDD puede utilizar un factor de ensanchamiento igual a 16 y transmisión multi-código, o bien un factor de ensanchamiento igual a 1 y transmisión multi-*slot* (un mismo usuario puede tener asignados varios *slots* en un mismo TTI). En TDD también resulta posible diferenciar varios usuarios dentro de un mismo TTI mediante una combinación de multiplexación por tiempo y por código.

Así pues, en el modo de operación FDD un canal físico queda definido por un código de canalización mientras que en el modo de operación TDD un canal físico queda definido por un código de canalización y por un *slot* dentro del TTI correspondiente.

A.3.2. Modos de transmisión en HSDPA

La tecnología HSDPA contempla la existencia de varios modos de transmisión cada uno de los cuales se basa en una combinación diferente de codificación de canal [209, 214] y modulación [210, 215]. Por ello, los modos de transmisión de HSDPA pueden denominarse MCSs (*Modulation and Coding Schemes*). Todos ellos se basan en modulaciones multinivel (QPSK y 16QAM). Inicialmente se consideró la posibilidad de emplear también modulaciones 8PSK y 64QAM aunque finalmente fueron descartadas por cuestiones de complejidad y rendimiento. Se emplea una codificación turbo de factor 1/3, aunque variando ciertos parámetros [166] se consiguen factores efectivos entre 1/4 y 3/4. La tabla A.3 muestra los parámetros más relevantes de cada modo de transmisión. Los valores mostrados se obtienen de la siguiente forma. En primer lugar, considerando una tasa de chip de 3.84 Mcps se obtiene un total de 2560 chips por *slot*. Puesto que en el canal HS-DSCH se emplean códigos con un factor de ensanchamiento igual a 16, en cada *slot* se transmiten 160 símbolos. El número de bits por símbolo dependerá del tipo de modulación empleada. En QPSK y en 16QAM se transmiten, respectivamente, 2 y 4 bits por símbolo modulado, lo cual se traduce en 320 bits por *slot* para QPSK y 640 bits por *slot* para 16QAM. Estos valores se refieren a la longitud de la secuencia obtenida tras el proceso de codificación. El factor del código empleado determina la cantidad de bits de información útil que son tomados a la entrada del codificador para dar como salida una secuencia de 320 bits por *slot* (QPSK) o de 640 bits por *slot* (16QAM). Para MCS-1 por ejemplo, tomando un bloque de 240 bits para ser transmitidos en un TTI de 2 ms, lo cual equivale a transmitir 80 bits de datos por *slot*, a la salida del codificador se obtienen los $80 \div 1/4 = 320$ bits por *slot* para QPSK. La cantidad de bits de datos por *slot* o por TTI se obtiene de la misma forma para el resto de modos de transmisión y, a partir del valor de dicho parámetro, se obtiene la tasa binaria en cada caso. Los valores mostrados en la tabla A.3 corresponden al máximo *throughput* que puede obtener un usuario que tiene asignado un único código de canalización. El máximo teórico en condiciones ideales se obtiene cuando la calidad del canal permite utilizar MCS-5 sin experimentar errores de transmisión y se asignan 15 códigos de canalización a un mismo usuario, dando lugar a 15×720 kbps $= 10.8$ Mbps. Las figuras A.10 y A.11 muestran las prestaciones de los diversos MCSs de HSDPA en función de la relación señal a interferencia (*Signal-to-Interference Ratio*, SIR).

Modo de transmisión	Tipo de modulación	Factor del código	Bits de datos por *slot*	Bits de datos por TTI de 2 ms	Tasa (kbps)
MCS-1	QPSK	1/4	80	240	120
MCS-2	QPSK	1/2	160	480	240
MCS-3	QPSK	3/4	240	720	360
MCS-4	16QAM	1/2	320	960	480
MCS-5	16QAM	3/4	480	1440	720

Tabla A.3: Modos de transmisión de HSDPA.

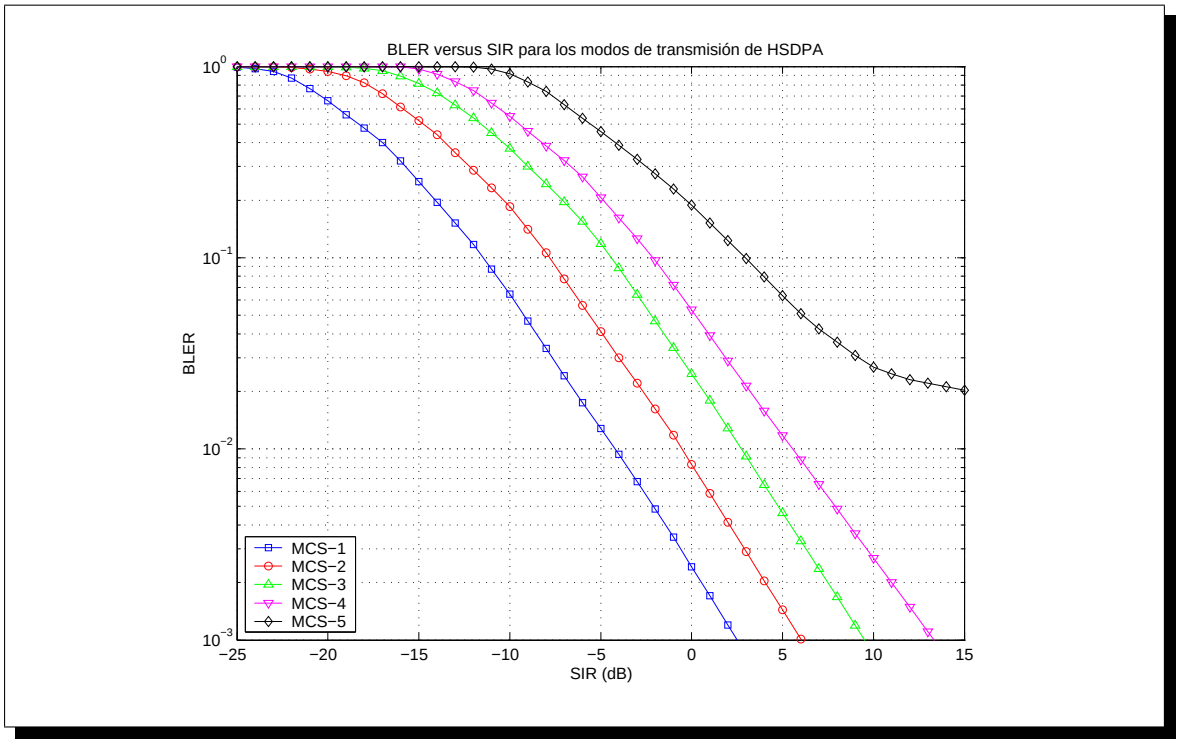


Figura A.10: BLER versus SIR para los modos de transmisión de HSDPA [219].

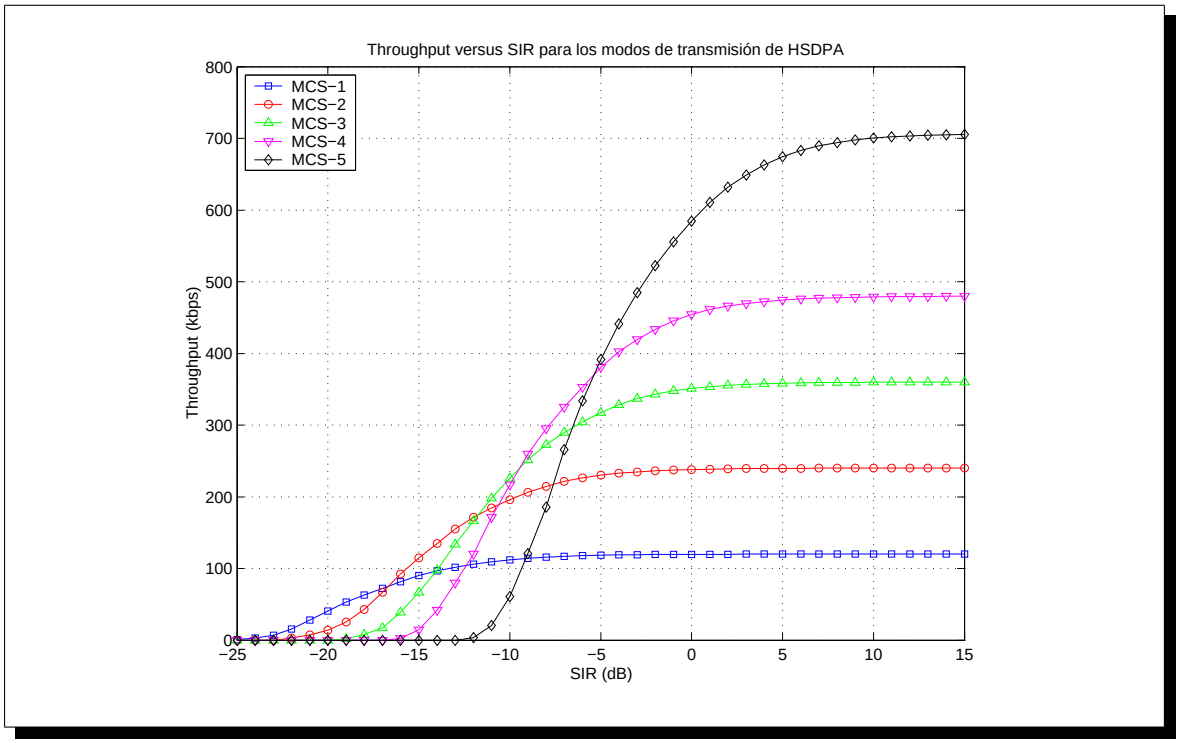


Figura A.11: *Throughput* versus SIR para los modos de transmisión de HSDPA.

La existencia de diversos modos de transmisión en HSDPA posibilita la aplicación de AMC (*Adaptive Modulation and Coding*). El principio de funcionamiento de AMC consiste en variar el formato de modulación y codificación en función de las variaciones del canal radio. Dicha variación se produce al seleccionar en cada momento el MCS más apropiado a la calidad experimentada en el enlace radio. Esta técnica se engloba dentro del conjunto de técnicas conocidas como *link adaptation*, en las cuales se modifica de forma dinámica ciertos parámetros de transmisión para compensar las variaciones de las condiciones de transmisión del canal y adaptar la transmisión de datos a través de la interfaz radio a la calidad experimentada en el enlace radio. En WCDMA *Release 99* la adaptación al enlace se realizaba variando el factor de ensanchamiento de los códigos de canalización asignados y ajustando la potencia de transmisión mediante algún algoritmo adecuado de control de potencia (para evitar los efectos de los desvanecimientos y mantener una comunicación relativamente estable). En HSDPA, la existencia de diversos modos de transmisión y la aplicación de AMC eliminan la necesidad del control de potencia (en HSDPA no existe control de potencia en el canal compartido HS-DSCH sino que se transmite a potencia constante) y de emplear códigos de canalización con factores de ensanchamiento variable (en HSDPA es constante e igual a 16). En este sentido, la aplicación de AMC reemplaza al procedimiento anterior de adaptación al enlace. En un sistema con AMC, los usuarios en posiciones favorables, es decir, cerca de la estación base, emplean MCSs con una modulación de orden elevado y un código de factor elevado, mientras que a los usuarios en posiciones desfavorables, es decir, lejos de la estación base, se les asignan MCSs con modulación de orden bajo y códigos con factores reducidos. Los beneficios de AMC son, principalmente, la disponibilidad de elevadas tasas de transmisión para los usuarios en posiciones favorables gracias a lo cual aumenta el *throughput* global de la celda y una baja variabilidad en la interferencia experimentada debido a que la adaptación al enlace se basa en variaciones de la codificación y modulación empleadas y no en variaciones de la potencia de transmisión.

En un sistema celular con tecnología HSDPA, la tasa de información se selecciona de forma *ruda* mediante AMC, que elige uno de los MCSs disponibles para transmisión. Mediante la aplicación de HARQ (*Hybrid Automatic Repeat reQuest*) se realiza un ajuste más fino de las tasas de codificación, mejorando así la precisión de la adaptación al enlace y la eficiencia de su utilización.

Concretamente, HARQ es una técnica que combina dos esquemas básicos de control de errores, uno de corrección de errores hacia delante (*Forward Error Correction*, FEC) y otro de corrección de errores hacia atrás (*Backward Error Correction*, BEC). El esquema FEC consiste en transmitir cierta información redundante junto con los datos de interés para posibilitar la detección y corrección de errores de transmisión en el extremo receptor. Este aspecto corresponde generalmente a la codificación y decodificación de canal y será el tratado en la presente sección. El esquema BEC suele implementarse mediante un protocolo ARQ donde el extremo receptor solicita la retransmisión de los bloques de datos erróneos (este aspecto será el tratado en la sección A.3.4). En un protocolo ARQ convencional, si un bloque de datos no puede decodificarse correctamente, directamente se descarta y se solicita su retransmisión. Cuando se recibe y decodifica el bloque retransmitido, no se utiliza información alguna proce-

dente de la transmisión previa de dicho bloque. En un esquema HARQ, a diferencia de un esquema ARQ convencional, el bloque de datos incorrecto no se descarta sino que se almacena en el *buffer* de recepción. Cuando se recibe la retransmisión de dicho bloque, se produce cierta combinación de la información retransmitida con la recibida en la transmisión anterior. Si la decodificación del bloque resultante vuelve a fallar, se vuelve a solicitar una nueva retransmisión, que volverá a ser combinada con las transmisiones anteriores del mismo bloque de datos antes de proceder a su decodificación. El proceso se repite hasta que el bloque de datos se decodifica correctamente o se alcanza un número máximo de retransmisiones, previamente establecido. La combinación de la transmisión inicial de un bloque de datos con sus posteriores retransmisiones aumenta la probabilidad de decodificar correctamente la información recibida en comparación con un esquema ARQ donde los bloques transmitidos se intentan decodificar de forma totalmente independiente. En [220] se distinguen tres tipos de HARQ (figura A.12):

Tipo I. En este tipo de HARQ se añade a los datos de interés un campo CRC para verificación de errores y se realiza una codificación FEC. En el receptor se realiza la decodificación FEC y se verifica la posible existencia de errores mediante el campo CRC. Si se detectan errores en el bloque recibido, se descarta dicho bloque y se solicita su retransmisión. El bloque retransmitido utilizará exactamente la misma codificación y será una replica exacta del bloque enviado en la transmisión inicial. Este tipo de HARQ realmente se basa en un esquema ARQ convencional.

Tipo II. Este tipo de HARQ también suele denominarse esquema de redundancia incremental (*Incremental Redundancy*, IR). En este tipo de HARQ los bloques recibidos con error no son descartados sino que son conservados en el *buffer* de recepción y combinados con las posteriores transmisiones asociadas a dicho bloque. En HARQ tipo II las retransmisiones consisten en información redundante adicional y distinta a la recibida en la transmisión inicial. Esta redundancia adicional se combina con la recibida en el bloque anterior, obteniéndose así una secuencia de bits con mayor cantidad de redundancia que aumenta la probabilidad de decodificar correctamente la información recibida. Cada retransmisión aporta únicamente información redundante (*Full IR*), distinta a la de las transmisiones anteriores. Por este motivo las retransmisiones no son autodecodificables y sólo pueden decodificarse después de combinarlas con las transmisiones previas.

Tipo III. Al igual que en HARQ tipo II, este tipo de HARQ también pertenece a los esquemas IR, lo cual implica que las transmisiones erróneas de un determinado bloque de datos no se descartan sino que se conservan y se combinan con posteriores retransmisiones. En HARQ tipo II las retransmisiones con redundancia adicional no son autodecodificables. En situaciones en las cuales el bloque transmitido inicialmente queda seriamente dañado por los errores de transmisión, resulta deseable que las futuras retransmisiones con redundancia adicional sean autodecodificables. En HARQ tipo III cada retransmisión es autodecodificable, lo cual significa que es posible recuperar la información transmitida a partir únicamente de la información retransmitida sin necesidad de combinarla con las transmisiones anteriores, si éstas han sido con una calidad pobre y las siguientes

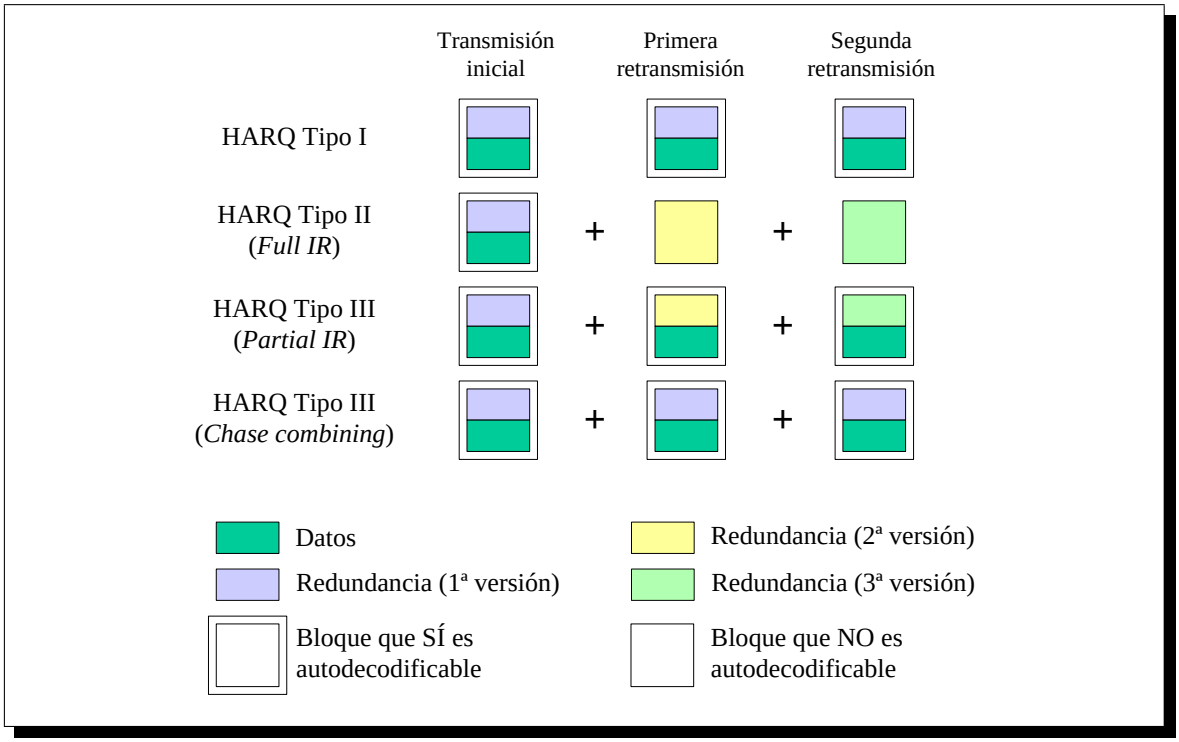


Figura A.12: Diferentes tipos de esquemas HARQ.

retransmisiones han experimentado una calidad suficientemente buena. Dentro de HARQ tipo III se distinguen dos subcasos. El primero de ellos se basa en el empleo de múltiples versiones de información redundante. En este caso cada bloque de datos se codifica empleando diferentes patrones de *puncturing*, dando lugar a bloques autodecodificables pero con diferentes versiones de redundancia. Si la primera versión del bloque se recibe con error, se reenvía la segunda y así sucesivamente. Resulta posible transmitir diferentes versiones de un mismo bloque de datos o repetir las disponibles cuando ya se han transmitido todas. Este caso es similar a HARQ tipo II (*Full IR*) pero ahora la información redundante ocupa sólo una parte de los bloques transmitidos, motivo por el cual suele denominarse *Partial IR*. El segundo subcaso de HARQ tipo III se basa en el empleo de una única versión de información redundante. El modo de operación es similar al de HARQ tipo I, con la diferencia de que las sucesivas retransmisiones de un mismo bloque de datos se combinan ponderadas por el valor de SIR experimentado en cada una de ellas según cierto método propuesto por Chase [221] y que recibe el nombre de *Chase combining*. En este tipo de esquema la repetición del código supone cierta ganancia por diversidad temporal, aunque existen algunas pérdidas por combinación de aproximadamente 0.2-0.3 dB por cada retransmisión [188].

Generalmente IR ofrece mejor rendimiento que *Chase combining* aunque posee mayor complejidad de implementación y no ofrece grandes resultados a menos que los errores de transmisión sean considerables. En cambio, el esquema *Chase combining* ofrece un rendimiento razonable con baja complejidad de implementación y bajo coste.

A.3.3. Transmisión de datos en HSDPA

Cuando la información procedente de las capas altas de la torre de protocolos llega a la capa MAC (*Medium Access Control*) [222], ésta se segmenta en unidades que reciben el nombre de *bloques de transporte* (*Transport Block*, TB). Un TB se define como la unidad básica de información intercambiada entre la capa MAC y la capa física [223]. En términos generales, la capa MAC transfiere los TBs a la capa física para su transmisión agrupados en conjuntos de varios TBs (*Transport Block Set*, TBS). Un TBS se define como un conjunto de TBs intercambiados entre la capa MAC y la capa física de forma simultánea utilizando el mismo canal de transporte [223]. En el caso del canal compartido HS-DSCH de HSDPA, un TBS contiene siempre un único TB. Evidentemente, la cantidad de bits de datos contenida en cada TB dependerá del modo de transporte que vaya a utilizarse para su transmisión, para el cual se define un determinado procedimiento de codificación, entrelazado, perforación o *puncturing*, modulación, etc., ya que el resultado de dicho proceso deberá ser siempre una secuencia que ha de ajustarse a la duración de un TTI. En HSDPA, por lo tanto, en un TTI se transmite un TBS, el cual contiene un único TB. El parámetro TTI se define como tiempo que transcurre entre la transferencia desde la capa MAC hacia la capa física de dos TBSs consecutivos [223]. En WCDMA *Release 99* puede ser igual a la duración de una trama (10 ms) o un múltiplo entero de dicho valor (20, 40 ó 80 ms). En HSDPA el valor del TTI es fijo y depende del modo de operación: 2 ms para HS-DSCH FDD, 10 ms para HS-DSCH TDD a 3.84 Mcps y 5 ms para HS-DSCH TDD a 1.28 Mcps.

Un aspecto importante en la transmisión de datos en HSDPA es que los usuarios reciben sus paquetes de datos a través de un canal de transporte compartido, canal que recibe el nombre de HS-DSCH. Los paquetes de los usuarios se colocan en colas de transmisión antes de ser transmitidos a sus correspondientes destinatarios a través de la interfaz radio. En el lado de la red existe una entidad denominada *scheduler* que, según cierto criterio, selecciona el usuario a servir en cada momento. En WCDMA *Release 99* esta entidad se encuentra en el RNC (*Radio Network Controler*) y decide las transmisiones y retransmisiones de información a nivel de bloque RLC. En HSDPA esta entidad ha sido desplazada hacia la interfaz radio para acercarla al terminal móvil, ubicándose en el Nodo B y decidiendo las transmisiones y retransmisiones de información a nivel de TBs, es decir, a nivel de unidades de datos de la capa MAC. Esta decisión, junto la reducción del TTI, pretende reducir los retardos de transmisión en la interfaz radio así como los tiempos de espera desde que se solicita la retransmisión de información hasta que ésta finalmente se produce. Con el fin de aprovechar al máximo estas posibilidades ofrecidas por HSDPA, es importante aplicar un criterio o algoritmo de *scheduling* eficiente, ya que dicha decisión repercute de forma directa sobre el rendimiento del sistema. En la literatura existen numerosos algoritmos de *scheduling*, aunque dentro de dicho conjunto puede identificarse dos extremos:

- *Round Robin*. Este algoritmo selecciona aquel usuario que, en el momento de la decisión, lleva más tiempo sin haber sido servido. El usuario seleccionado accede al canal HS-DSCH durante un intervalo de tiempo máximo predefinido que se

denomina *cuanto*. Si al agotarse el cuanto temporal el usuario no ha transmitido toda la información disponible, vuelve a ponerse en la cola de espera.

- *Maximum CIR*. Este algoritmo selecciona aquel usuario que, en el momento de la decisión, experimenta la mejor calidad instantánea en el enlace radio expresada en términos de la relación portadora a interferencia (CIR).

El algoritmo *Round Robin* es el más justo y equitativo en la medida en que los diferentes usuarios del sistema tienen las mismas posibilidades de acceso al recurso de transmisión, aunque este criterio generalmente no ofrece los mejores resultados posibles al permitir que cualquier usuario pueda transmitir independientemente de la calidad de transmisión experimentada. El algoritmo *Maximum CIR* ha demostrado resultar óptimo a la hora de obtener el mejor rendimiento global a nivel de sistema, aunque en contraposición con el algoritmo *Round Robin* resulta mucho menos equitativo. En efecto, un usuario cercano a la estación base que se encuentre en condiciones favorables puede acaparar por completo el uso del recurso de transmisión, impidiendo el acceso por parte de otros usuarios más alejados y, por tanto, en condiciones menos favorables. Entre estos dos extremos existen numerosos algoritmos de *scheduling* en la literatura y el criterio seleccionado es una cuestión específica del fabricante de los equipos.

A.3.4. Protocolo de retransmisión en HSDPA

En la sección A.3.2 se señaló el empleo en HSDPA de la técnica HARQ que combina aspectos relativos tanto a la codificación y decodificación de canal de la información transmitida como al protocolo de retransmisión empleado. El primer aspecto fue descrito en la sección A.3.2. El segundo de ellos se describe en la presente sección.

El protocolo de retransmisión de datos seleccionado para la operación de HARQ en HSDPA es el conocido como parada-y-espera (*Stop-and-Wait*, SAW). En este protocolo el transmisor persiste en la transmisión de un TB determinado hasta que éste se recibe de forma correcta antes de iniciar la transmisión del siguiente TB. El correcto funcionamiento del protocolo queda asegurado mediante un simple bit utilizado para indicar el número de secuencia del TB actual o del siguiente, dando lugar a una carga de control mínima. La carga asociada a las confirmaciones es también mínima, ya que la recepción correcta o incorrecta del TB actual puede indicarse también mediante un único bit. Además, puesto que en todo momento sólo existe un único TB en transmisión, los requisitos de memoria en el terminal móvil son también mínimos. No obstante, existe una importante desventaja en este protocolo: las confirmaciones enviadas por el receptor no llegan al receptor de forma inmediata, lo cual obliga al extremo transmisor a permanecer en espera hasta que recibe la confirmación del receptor. Este es un problema bien conocido del protocolo SAW, que provoca que el canal de transmisión esté ocioso durante ciertos intervalos de tiempo en los cuales la capacidad de transmisión se desperdicia. En un canal de transmisión ranurado, como es el caso del canal HS-DSCH, el retardo de confirmación supone al menos la mitad del intervalo de transmisión, lo cual significa desperdiciar un intervalo de transmisión por cada transmisión,

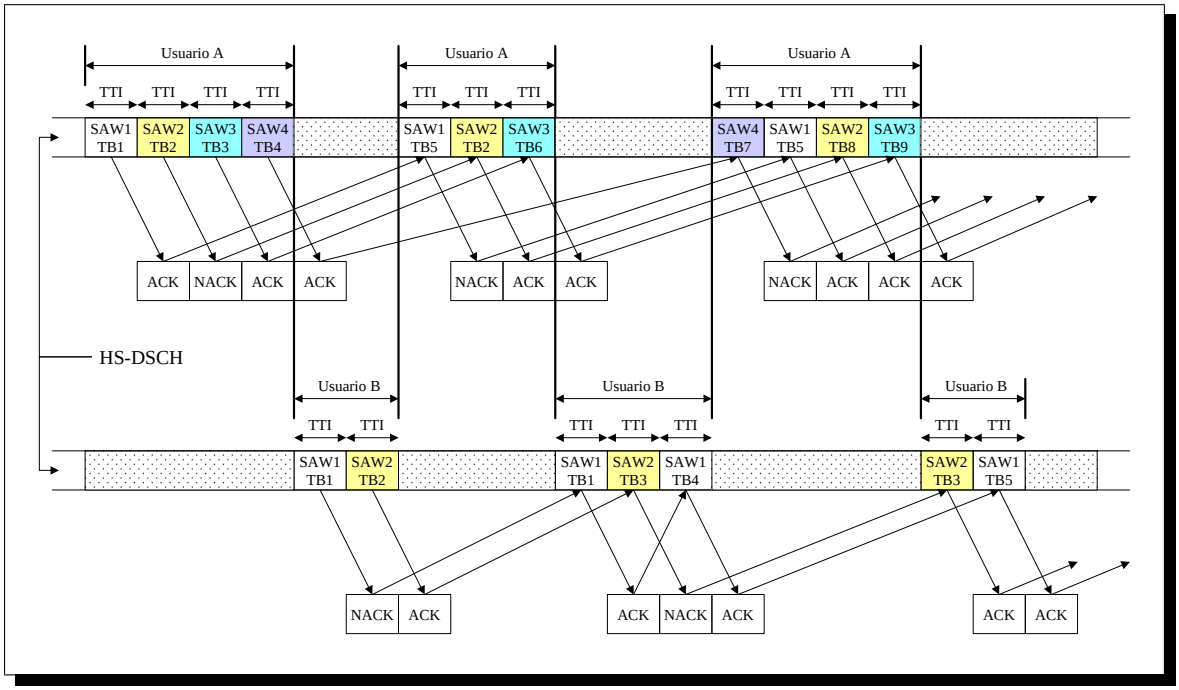


Figura A.13: Retransmisión de datos en HSDPA.

es decir, la mitad de la capacidad de transmisión. La solución consiste en emplear N procesos SAW en paralelo (hasta 8 como máximo [218]), de manera que cuando uno de ellos espera confirmación del extremo receptor, otro accede al canal y realiza una transmisión. De esta forma no se desperdicia capacidad de transmisión alguna, aunque en este caso el *buffer* receptor debe tener capacidad de almacenamiento suficiente como para poder almacenar N bloques diferentes.

La figura A.13 muestra un ejemplo del funcionamiento del protocolo SAW con N procesos simultáneos, en el cual dos usuarios A y B transmiten a través del canal HS-DSCH. Las transmisiones de ambos usuarios se muestran de forma separada para facilitar la comprensión, aunque ambas transmisiones realmente tienen lugar sobre el mismo canal HS-DSCH. Los intervalos de tiempo mostrados durante los cuales cada usuario accede al canal HS-DSCH son arbitrarios (en el ámbito práctico dependerán de la política de *scheduling* empleada). El usuario A transmite haciendo uso de $N = 4$ procesos SAW simultáneos mientras que el usuario B lo hace utilizando $N = 2$ procesos. Puede observarse que, para un mismo usuario, los procesos SAW se alternan de forma sucesiva cada vez que el usuario tiene acceso al canal HS-DSCH. Cada vez que un determinado proceso debe transmitir en el TTI asignado, transmite un nuevo TB si el transmitido anteriormente llegó correctamente al receptor (ACK) o retransmite el TB transmitido en la ocasión anterior si éste llegó con error al receptor (NACK). Puesto que las retransmisiones se realizan empleando el mismo modo de transmisión que en la transmisión inicial, el número de retransmisiones permitidas para un mismo TB suele limitarse a un valor prefijado para evitar que las retransmisiones de TBs erróneos se prolonguen de manera indefinida cuando las condiciones de transmisión radio no son favorables. Una vez alcanzado dicho límite, se puede cambiar de modo de transmisión.

BIBLIOGRAFÍA

[1] A. Kucar, "Mobile Radio: An Overview", *IEEE Communications Magazine*, pp. 72–85, Noviembre 1991.

[2] M. Mouly, M.-B. Pautet, *The GSM System for Mobile Communications*, Palaiseau, 1992.

[3] T. Halonen, J. Romero, J. Melero, *GSM, GPRS and EDGE Performance: Evolution Towards 3G/UMTS*, Wiley, 2ª edición, 2002.

[4] J. De Vriendt, P. Lainé, C. Lerouge, X. Xu, "Mobile Network Evolution: A Revolution on the Move", *IEEE Communications Magazine*, vol. 40, nº 4, pp. 104–111, Abril 2002.

[5] W. Mohr, W. Konhäuser, "Access Network Evolution Beyond Third Generation Mobile Communications", *IEEE Communications Magazine*, vol. 38, nº 12, pp. 122–133, Diciembre 2000.

[6] A. Capone, L. Musumeci, "Performance Evaluations of High Speed Circuit Data Services in GSM Systems", *IEEE 47th Vehicular Technology Conference*, 4-7 Mayo 1997, vol. 3, pp. 1684–1688.

[7] D. Calin, D. Zeghlache, "Performance Analysis of High Speed Circuit Switched Data (HSCSD) over GSM", *IEEE International Conference on Communications*, 7-11 Junio 1998, vol. 3, pp. 1586–1590.

[8] J. H. Sarker, S. J. Halme, M. Rinne, "Performance Analysis of GSM Traffic Channel Capacity With(out) High Speed Circuit Switched Data", *52nd Vehicular Technology Conference*, 24-28 Septiembre 2000, vol. 4, pp. 1603–1609.

[9] C. Bettstetter, H.-J. Vögel, J. Eberspächer, "GSM Phase 2+ General Packet Radio Service GPRS: Architecture, Protocols, and Air Inter-face", *IEEE Communications Surveys*, vol. 2, nº 3, pp. 2–14, 3er. Cuatrimestre 1999.

[10] G. Brasche, B. Walke, "Concepts, Services, and Protocols of the New GSM Phase 2+ General Packet Radio Service", *IEEE Communications Magazine*, vol. 35, nº 8, pp. 94–104, Agosto 1997.

[11] J. Cai, D. J. Goodman, "General Packet Radio Service in GSM", *IEEE Communications Magazine*, vol. 35, nº 10, pp. 122–131, Octubre 1997.

[12] A. Furuskär, M. Frodigh, H. Olofsson, J. Sköld, "System Performance of EDGE, a Proposal for Enhanced Data Rates in Existing Digital Cellular Systems", *48th IEEE Vehicular Technology Conference*, Ottawa (Canada), 18-21 Mayo 1998, vol. 2, pp. 1284–1289.

[13] A. Furuskär, S. Mazur, F. Müller, H. Olofsson, "EDGE: Enhanced Data Rates for GSM and TDMA/136 Evolution", *IEEE Personal Communications*, vol. 6, nº 3, pp. 56–66, Junio 1999.

[14] A. Furuskär, M. Höök, S. Jäverbring, H. Olofsson, J. Sköld, "Capacity Evaluation of the EDGE Concept for Enhanced Data Rates in GSM and TDMA/136", *49th IEEE Vehicular Technology Conference*, Houston (Estados Unidos), 16-20 Mayo 1999, vol. 2, pp. 1648–1652.

[15] E. Dahlman, B. Gudmundson, M. Nilsson, A. Skold, "UMTS/IMT-2000 Based on Wideband CDMA", *IEEE Communications Magazine*, vol. 36, nº 9, pp. 70–80, Septiembre 1998.

[16] B. Melis, G. Romano, "UMTS W-CDMA: Evaluation of Radio Performance by means of Link Level Simulations", *IEEE Personal Communications*, vol. 7, nº 3, pp. 42–49, Junio 2000.

[17] IEEE 802.15 Working Group for WPAN, URL: http://www.ieee802.org/15/ y http://grouper.ieee.org/groups/802/15/.

[18] IEEE 802 Wireless World, URL: http://www.802wirelessworld.com/index.jsp.

[19] IEEE Wireless Standards Zone, URL: http://standards.ieee.org/wireless/.

[20] The Official Bluetooth Membership Site, URL: https://www.bluetooth.org/.

[21] IEEE 802.11 Wireless Local Area Networks, URL: http://www.ieee802.org/11/ y http://grouper.ieee.org/groups/802/11/.

[22] ETSI Radio Competence Centre, HIgh PErformance Radio Local Area Network (HIPERLAN), URL: http://portal.etsi.org/radio/HiperLAN/HiperLAN.asp.

[23] J. Khun-Jush, P. Schramm, G. Malmgren, J. Torsner, "HiperLAN2: Broadband Wireless Communications at 5 GHz", *IEEE Communications Magazine*, vol. 40, nº 6, pp. 130–136, Junio 2002.

[24] J. Khun-Jush, G. Malmgren, P. Schramm, J. Torsner, "Overview and Performance of HIPERLAN Type 2-a Standard for Broadband Wireless Communications", *IEEE 51st Vehicular Technology Conference Proceedings*, Tokyo (Japón), 15-18 Mayo 2000, vol. 1, pp. 112–117.

[25] IEEE 802.16 Working Group on Broadband Wireless Access Standards, URL: http://www.ieee802.org/16/ y http://grouper.ieee.org/groups/802/16/.

[26] ETSI Radio Competence Centre, HIgh PERformance Metropolitan Area Network (HiperMAN), URL: http://portal.etsi.org/radio/HiperMAN/HiperMAN.asp.

[27] WiMAX Forum, URL: http://www.wimaxforum.org/home.

[28] Digital Video Broadcasting (DVB), URL: http://www.dvb.org/.

[29] ETSI Radio Competence Centre, Digital Video Broadcasting (DVB), URL: http://portal.etsi.org/radio/DigitalVideoBroadcasting/DVB.asp.

[30] ETSI Radio Competence Centre, Digital Audio Broadcasting (DAB), URL: http://portal.etsi.org/radio/DigitalAudioBroadcasting/DAB.asp.

[31] M. Progler, C. Evci, M. Umehira, "Air Interface Access Schemes for Broadband Mobile Systems", *IEEE Communications Magazine*, vol. 37, nº 9, pp. 106–115, Septiembre 1999.

[32] T. Tjelta, A. Nordbotten, M. Annoni, E. Scarrone, S. Bizzarri, L. Tokarchuk, J. Bigham, C. Adams, K. Craig, M. Dinis, "Future Broadband Radio Access Systems for Integrated Services with Flexible Resource Management", *IEEE Communications Magazine*, vol. 39, nº 8, pp. 56–63, Agosto 2001.

[33] U. Varshney, R. Vetter, "Emerging Mobile and Wireless Networks", *Communications of the ACM*, vol. 43, nº 6, pp. 73–81, Junio 2000.

[34] F. Sestini, "European R&D Beyond Third Generation Mobile Communication Systems", *IEEE International Conference on Personal Wireless Communications*, 17-20 Diciembre 2000, pp. 563–567.

[35] F. Sestini, J. S. da Silva, J. Fernandes, "Expanding the Wireless Universe: EU Research on the Move", *IEEE Communications Magazine*, vol. 40, nº 10, pp. 132–140, Octubre 2002.

[36] C. Evci, U. Barth, P. Sehier, R. Sigle, "The Path to Beyond 3G Systems: Strategic and Technological Challenges", *4th International Conference on 3G Mobile Communication Technologies*, 25-27 Junio 2003, pp. 299–303.

[37] H.-B. Lim, "Beyond 3G: Issues and Challenges", *IEEE Potentials*, vol. 21, nº 4, pp. 18–23, Octubre - Noviembre 2002.

[38] M. Flament, F. Gessler, F. Lagergren, O. Queseth, R. Stridh, M. Unbehaun, J. Wu, J. Zander, "An Approach to 4th Generation Wireless Infrastructures - Scenarios and Key Research Issues", *IEEE 49th Vehicular Technology Conference*, 16-20 Mayo 1999, vol. 2, pp. 1742–1746.

[39] A. Bria, F. Gessler, O. Queseth, R. Stridh, M. Unbehaun, J. Wu, J. Zander, M. Flament, "4th-Generation Wireless Infrastructures: Scenarios and Research Challenges", *IEEE Personal Communications*, vol. 8, nº 6, pp. 25–31, Diciembre 2001.

[40] I. Ganchev, M. O'Droma, H. Chaouchi, I. Armuelles, M. Siebert, N. Houssos, "Requirements for an Integrated System and Service 4G Architecture", *IEEE 59th Vehicular Technology Conference*, 17-19 Mayo 2004, vol. 5, pp. 3029–3034.

[41] M. Hata, "Fourth Generation Mobile Communication Systems Beyond IMT-2000 Communications", *Fifth Asia-Pacific Conference on Communications y Fourth Optoelectronics and Communications Conference*, Beijing (China), 18-22 Octubre 1999, vol. 1, pp. 765–767.

[42] K. Murota, "Mobile Communications Trends in Japan and DoCoMo's Activities towards 21st Century", *4th ACTS Mobile Communications Summit*, Sorrento (Italia), 8-11 Junio 1999.

[43] J. M. Pereira, "Fourth Generation: Now, It Is Personal!", *The 11th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 18-21 Septiembre 2000, vol. 2, pp. 1009–1016.

[44] A. Doufexi, E. Tameh, A. Nix, S. Armour, A. Molina, "Hotspot Wireless LANs to Enhance the Performance of 3G and Beyond Cellular Networks", *IEEE Communications Magazine*, vol. 41, nº 7, pp. 58–65, Julio 2003.

[45] H. Honkasalo, K. Pehkonen, M. T. Niemi, A. T. Leino, "WCDMA and WLAN for 3G and Beyond", *IEEE Wireless Communications*, vol. 9, nº 2, pp. 14–18, Abril 2002.

[46] A. K. Salkintzis, "Interworking Techniques and Architectures for WLAN/3G Integration toward 4G Mobile Data Networks", *IEEE Wireless Communications*, vol. 11, nº 3, pp. 50–61, Junio 2004.

[47] A. K. Salkintzis, "WLAN/3G Interworking Architectures for Next Generation Hybrid Data Networks", *IEEE International Conference on Communications*, 20-24 Junio 2004, vol. 7, pp. 3984–3988.

[48] A. K. Salkintzis, "Interworking between WLANs and Third-Generation Cellular Data Networks", *The 57th IEEE Semiannual Vehicular Technology Conference*, 22-25 Abril 2003, vol. 3, pp. 1802–1806.

[49] K. Ahmavaara, H. Haverinen, R. Pichna, "Interworking Architecture between 3GPP and WLAN Systems", *IEEE Communications Magazine*, vol. 41, nº 11, pp. 74–81, Noviembre 2003.

[50] A. K. Salkintzis, C. Fors, R. Pazhyannur, "WLAN-GPRS Integration for Next-Generation Mobile Data Networks", *IEEE Wireless Communications*, vol. 9, nº 5, pp. 112–124, Octubre 2002.

[51] M. Buddhikot, G. Chandranmenon, S. Han, Y.-W. Lee, S. Miller, L. Salgarelli, "Integration of 802.11 and Third-Generation Wireless Data Networks", *Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies*, 30 Marzo - 3 Abril 2003, vol. 1, pp. 503–512.

[52] M. M. Buddhikot, G. Chandranmenon, S. Han, Y.-W. Lee, S. Miller, L. Salgarelli, "Design and Implementation of a WLAN/CDMA2000 Interworking Architecture", *IEEE Communications Magazine*, vol. 41, nº 11, pp. 90–100, Noviembre 2003.

[53] D. Kouis, P. Demestichas, V. Stavroulaki, G. Koundourakis, N. Koutsouris, L. Papadopoulou, N. Mitrou, "System for Enhanced Network Management Jointly Exploiting WLANs and Other Wireless Network Infrastructures", *IEE Proceedings-Communications*, vol. 151, nº 5, pp. 514–520, 24 Octubre 2004.

[54] W. Zhuang, Y.-S. Gan, K.-J. Loh, K.-C. Chua, "Policy-Based QoS-Management Architecture in an Integrated UMTS and WLAN Environment", *IEEE Communications Magazine*, vol. 41, nº 11, pp. 118–125, Noviembre 2003.

[55] J. Luo, R. Mukerjee, M. Dillinger, E. Mohyeldin, E. Schulz, "Investigation of Radio Resource Scheduling in WLANs Coupled with 3G Cellular Network", *IEEE Communications Magazine*, vol. 41, nº 6, pp. 108–115, Junio 2003.

[56] R. Pintenet, J. Gosteau, T. Al-Gizawi, F. Lazarakis, K. Peppas, A. Alexiou, "Evaluation of Interoperability Mechanisms for Coexisting HSDPA and WLAN Enhanced with MTMR Techniques", *IEEE 58th Vehicular Technology Conference*, 6-9 Octubre 2003, vol. 3, pp. 1807–1811.

[57] O. Benali, K. El-Khazen, D. Garrec, M. Guiraudou, G. Martínez, "A Framework for an Evolutionary Path Toward 4G by means of Cooperation of Networks", *IEEE Communications Magazine*, vol. 42, nº 5, pp. 82–89, Mayo 2004.

[58] J. Huschke, W. Rave, T. Kohler, "Downlink Capacity of UTRAN Reusing Frequencies of a DVB-T Network with Negligible Influence on DVB-T Performance", *IEEE 55th Vehicular Technology Conference*, 6-9 Mayo 2002, vol. 3, pp. 1579–1583.

[59] W. Klingenberg, A. Neutel, "MEMO: A Hybrid DAB/GSM Communication System for Mobile Interactive Multimedia Services", *European Conference on Multimedia Applications, Services and Techniques*, Berlín (Alemania), 1998, vol. 1425 of *Lecture Notes in Computer Science*, pp. 493–503.

[60] S. Uskela, "Key Concepts for Evolution toward Beyond 3G Networks", *IEEE Wireless Communications*, vol. 10, nº 1, pp. 43–48, Febrero 2003.

[61] R. Keller, T. Lohmar, R. Tönjes, J. Thielecke, "Convergence of Cellular and Broadcast Networks from a Multi-Radio Perspective", *IEEE Personal Communications*, vol. 8, nº 2, pp. 51–56, Abril 2001.

[62] C. Politis, T. Oda, S. Dixit, A. Schieder, H.-Y. Lach, M. I. Smirnov, S. Uskela, R. Tafazolli, "Cooperative Networks for the Future Wireless World", *IEEE Communications Magazine*, vol. 42, nº 9, pp. 70–79, Septiembre 2004.

[63] L. Becchetti, F. D. Priscoli, T. Inzerilli, P. Mähönen, L. Muñoz, "Enhancing IP Service Provision over Heterogeneous Wireless Networks: A Path Toward 4G", *IEEE Communications Magazine*, vol. 39, nº 8, pp. 74–81, Agosto 2001.

[64] S. Frattasi, R. L. Olsen, M. de Sanctis, F. H. P. Fitzek, R. Prasad, "Heterogeneous Services and Architectures for Next-Generation Wireless Networks", *2nd International Symposium on Wireless Communication Systems*, Siena (Italia), 5-9 Septiembre 2005, pp. 213–217.

[65] R. E. Schuh, P. Eneroth, P. Karlsson, "Multi-Standard Mobile Terminals", *IST Mobile & Wireless Telecommunications Summit*, Thessaloniki (Grecia), 16-19 Junio 2002, pp. 174–178.

[66] M. Mehta, N. Drew, G. Vardoulias, N. Greco, C. Niedermeier, "Reconfigurable Terminals: An Overview of Architectural Solutions", *IEEE Communications Magazine*, vol. 39, nº 8, pp. 82–89, Agosto 2001.

[67] J. Palicot, C. Roland, "A New Concept for Wireless Reconfigurable Receivers", *IEEE Communications Magazine*, vol. 41, nº 7, pp. 124–133, Julio 2003.

[68] P. Demestichas, G. Vivier, K. El Khazen, M. Theologou, "Evolution in Wireless Systems Management Concepts: From Composite Radio Environments to Reconfigurability", *IEEE Communications Magazine*, vol. 42, nº 5, pp. 90–98, Mayo 2004.

[69] P. Demestichas, L. Papadopoulou, V. Stavroulaki, M. Theologou, G. Vivier, G. Martinez, F. Galliano, "Wireless Beyond 3G: Managing Services and Network Resources", *Computer*, vol. 35, nº 8, pp. 80–82, Agosto 2002.

[70] P. Demestichas, N. Koutsouris, G. Koundourakis, K. Tsagkaris, A. Oikonomou, V. Stavroulaki, L. Papadopoulou, M. E. Theologou, G. Vivier, K. El Khazen, "Management of Networks and Services in a Composite Radio Context", *IEEE Wireless Communications*, vol. 10, nº 4, pp. 44–51, Agosto 2003.

[71] V. Stavroulaki, P. Demestichas, L. Papadopoulou, M. Theologou, "Traffic Distribution in Composite Radio Environments", *The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 15-18 Septiembre 2002, vol. 5, pp. 2269–2273.

[72] P. Demestichas, V. Stavroulaki, L. Papadopoulou, A. Vasilakos, M. Theologou, "Service Configuration and Traffic Distribution in Composite Radio Environments", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 34, nº 1, pp. 69–81, Febrero 2004.

[73] S. Y. Hui, K. H. Yeung, "Challenges in the Migration to 4G Mobile Systems", *IEEE Communications Magazine*, vol. 41, nº 12, pp. 54–59, Diciembre 2003.

[74] R. Walsh, L. Xu, T. Paila, "Hybrid Networks - A Step Beyond 3G", *3rd International Symposium on Wireless Personal Multimedia Communications*, Bangkok (Tailandia), Noviembre 2000, pp. 109–114.

[75] P. Beming, M. Cramby, N. Johansson, J. Lundsjö, G. Malmgren, J. Sachs, "Beyond 3G Radio Access Network Reference Architecture", *IEEE 59th Vehicular Technology Conference*, 17-19 Mayo 2004, vol. 4, pp. 2047–2051.

[76] G. Wu, M. Mizuno, P. J. M. Havinga, "MIRAI Architecture for Heterogeneous Network", *IEEE Communications Magazine*, vol. 40, nº 2, pp. 126–134, Febrero 2002.

[77] U. Varshney, R. Jain, "Issues in Emerging 4G Wireless Networks", *Computer*, vol. 34, nº 6, pp. 94–96, Junio 2001.

[78] X. Gao, G. Wu, T. Miki, "End-to-End QoS Provisioning in Mobile Heterogeneous Networks", *IEEE Wireless Communications*, vol. 11, nº 3, pp. 24–34, Junio 2004.

[79] X. Gao, G. Wu, T. Miki, "End-to-End QoS Provisioning in Mobile Heterogeneous Networks", *IEEE Wireless Communications and Networking*, 16-20 Marzo 2003, vol. 2, pp. 1361–1366.

[80] X. Gao, G. Wu, T. Miki, "QoS Framework for Mobile Heterogeneous Networks", *IEEE International Conference on Communications*, 11-15 Mayo 2003, vol. 2, pp. 933–937.

[81] D. Mandato, A. Kassler, T. R. Valladares, G. Neureiter, "Handling End-to-End QoS in Mobile Heterogeneous Networking Environments", *12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 30 Septiembre - 3 Octubre 2001, vol. 1, pp. C–49 – C–54.

[82] A. Kassler, T. Guenkova-Luy, D. Mandato, P. Schoo, I. Armuelles, T. Robles, P. Ruiz, A. Bascuñana, "Enabling Mobile Heterogeneous Networking Environments with End-to-End User Perceived QoS - The BRAIN Vision and the MIND Approach", *European Wireless*, Florencia (Italia), 25-28 Febrero 2002.

[83] R. Jain, U. Varshney, "Supporting Quality of Service in Multiple Heterogeneous Wireless Networks", *IEEE 55th Vehicular Technology Conference*, 6-9 Mayo 2002, vol. 2, pp. 952–956.

[84] J. Manner, L. Burness, E. Hepworth, A. López, E. Mitjana, "Provision of QoS in Heterogeneous Wireless IP Access Networks", *The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 15-18 Septiembre 2002, vol. 2, pp. 530–534.

[85] R. Ferrús, A. Gelonch, O. Sallent, J. Pérez-Romero, N. Nafisi, M. Dohler, "A Feasible Approach for QoS Management in Coordinated Heterogeneous Radio Access Networks", *24th IEEE International Performance, Computing, and Communications Conference*, 7-9 Abril 2005, pp. 607–614.

[86] L. Jorguseski, R. Prasad, "Overview of Radio Resource Management (RRM) Issues in Multi-Radio Access Systems", *7th European Conference on Wireless Technology*, 2004, pp. 97–100.

[87] P. Leaves, K. Moessner, R. Tafazolli, D. Grandblaise, D. Bourse, R. Tönjes, M. Breveglieri, "Dynamic Spectrum Allocation in Composite Reconfigurable Wireless Networks", *IEEE Communications Magazine*, vol. 42, nº 5, pp. 72–81, Mayo 2004.

[88] P. Leaves, S. Ghaheri-Niri, R. Tafazolli, L. Christodoulides, T. Sammut, W. Stahl, J. Huschke, "Dynamic Spectrum Allocation in a Multi-Radio Environment: Concept and Algorithm", *Second International Conference on 3G Mobile Communication Technologies*, 26-28 Mayo 2001, pp. 53–57.

[89] P. Leaves, S. Ghaheri-Niri, R. Tafazolli, J. Huschke, "Dynamic Spectrum Allocation in Hybrid Networks with Imperfect Load Prediction", *Third International Conference on 3G Mobile Communication Technologies*, 8-10 Mayo 2002, pp. 444–448.

[90] P. Leaves, J. Huschke, R. Tafazolli, "A Summary of Dynamic Spectrum Allocation Results from DRiVE", *IST Mobile and Wireless Telecommunications Summit*, Thessaloniki (Grecia), 16-19 Junio 2002, pp. 245–250.

[91] P. Leaves, R. Tafazolli, "A Time-Adaptive Dynamic Spectrum Allocation Scheme for a Converged Cellular and Broadcast System", *IEE Getting the Most Out of the Radio Spectrum Conference*, Londres (Reino Unido), 24-25 Octubre 2002, pp. 18/1–18/5.

[92] P. Leaves, M. Breveglieri, C. Hamacher, D. Grandblaise, F. Migneret, K. Moessner, D. Bourse, "Dynamic Spectrum Allocation and System Coexistence in Reconfigurable Multi-Radio Networks", *IST Mobile Summit*, Aveiro (Portugal), 16-18 Junio 2003.

[93] D. Grandblaise, K. Moessner, P. Leaves, D. Bourse, "Dynamic Spectrum Allocation (DSA): From Concept to Implementation", *International Workshop on Reconfigurable Terminals and Supporting Networks, Joint IST SCOUT Project and SDR Forum Workshop*, París (Francia), 16 Septiembre 2003.

[94] P. Leaves, M. Breveglieri, D. Grandblaise, C. Hamacher, F. Migneret, "Dynamic Spectrum Allocation for Systems Beyond 3G", *B3G Report, A Publication by Systems Beyond 3G Cluster*, 31 Octubre 2003, pp. 96–103.

[95] M. K. Pereirasamy, J. Luo, M. Dillinger, C. Hartmann, "An Approach for Inter-operator Spectrum Sharing for 3G Systems and Beyond", *15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 5-8 Septiembre 2004, vol. 3, pp. 1952–1956.

[96] P. Krishnamurthy, K. Pahlavan, A. Zahedi, J. Vallström, J. Talvitie, R. Pichna, M. Ylianttila, "Handoff in 3G Non-Homogenous Mobile Data Networks", *European Wireless*, Amsterdam (Holanda), Octubre 1998.

[97] H. J. Wang, R. H. Katz, J. Giese, "Policy-Enabled Handoffs Across Heterogeneous Wireless Networks", *Second IEEE Workshop on Mobile Computing Systems and Applications*, 25-26 Febrero 1999, pp. 51–60.

[98] K. Pahlavan, P. Krishnamurthy, A. Hatami, M. Ylianttila, J.-P. Makela, R. Pichna, J. Vallström, "Handoff in Hybrid Mobile Data Networks", *IEEE Personal Communications*, vol. 7, nº 2, pp. 34–47, Abril 2000.

[99] P. Demestichas, V. Stavroulaki, "Issues in Introducing Resource Brokerage Functionality in B3G Composite Radio Environments", *IEEE Wireless Communications*, vol. 11, nº 5, pp. 32–40, Octubre 2004.

[100] P. Demestichas, A. Oikonomou, K. Floros, L. Papadopoulou, V. Stavroulaki, M. Theologou, G. Martinez, G. Vivier, F. Galliano, E. S. de la Fuente Gallego, D. Zeglache, P. Kuonen, C. Caragiuli, "Enhanced Service Availability and Resource Utilisation by means of Composite Radio Systems Management", *12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 30 Septiembre - 3 Octubre 2001, vol. 2, pp. G–141 – G–145.

[101] S. Sumaryo, E. Hepworth, D. J. Higgins, M. Siebert, "A Radio Resource Management Architecture for a Beyond-3G Network", *International Workshop on Mobile IP-based Network Developments*, Budapest University of Technology and Economics (Hungría), Noviembre 2002.

[102] J. Pérez-Romero, O. Sallent, R. Agustí, "On Evaluating Beyond 3G Radio Access Networks: Architectures, Approaches and Tools", *IEEE 61st Vehicular Technology Conference*, 30 Abril - 1 Mayo 2005, vol. 5, pp. 2964–2968.

[103] G. D. Morley, W. D. Grover, "Strategies to Maximize Carried Traffic in Dual-Mode Cellular Systems", *IEEE Transactions on Vehicular Technology*, vol. 49, nº 2, pp. 357–366, Marzo 2000.

[104] R. Ramesh, K. Balachandran, "Traffic Analysis of Partially Overlaid AMPS/IS-136 Systems", *The Ninth IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 8-11 Septiembre 1998, vol. 1, pp. 288–292.

[105] H. Tawfik, "Traffic Advantages of Dual-Mode North American Cellular System", *IEEE 43rd Vehicular Technology Conference*, 18-20 Mayo 1993, pp. 202–205.

[106] H. Tawfik, "A New Channel Assignment Algorithm for Dual-Mode Cellular Systems", *Proceedings of the Sixteenth National Radio Science Conference*, 23-25 Febrero 1999, pp. C10/1–C10/8.

[107] T. W. Wong, V. K. Prabhu, "Multi-Mode/Multi-Carrier Resource Management in CDMA/AMPS Deployment", *IEEE 49th Vehicular Technology Conference*, vol. 3, pp. 1861–1865, 16-20 Mayo 1999.

[108] J. Kalliokulju, P. Meche, M. J. Rinne, J. Vallström, P. Varshney, S.-G. Häggman, "Radio Access Selection for Multistandard Terminals", *IEEE Communications Magazine*, vol. 39, nº 10, pp. 116–124, Octubre 2001.

[109] J. Pérez-Romero, O. Sallent, R. Agustí, "Policy-Based Initial RAT Selection Algorithms in Heterogeneous Networks", *The 7th IFIP International Conference on Mobile and Wireless Communications Networks*, Marrakech (Marruecos), 19-21 Septiembre 2005.

[110] J. Pérez-Romero, O. Sallent, R. Agustí, "On the Capacity Degradation in W-CDMA Uplink/Downlink Due to Indoor Traffic", *IEEE 60th Vehicular Technology Conference*, Los Ángeles (Estados Unidos), 26-29 Septiembre 2004, vol. 2, pp. 856–859.

[111] X. Gelabert, J. Pérez-Romero, O. Sallent, R. Agustí, "On the Impact of Multi-Mode Terminals in Heterogeneous Wireless Access Networks", *2nd International Symposium on Wireless Communication Systems*, Siena (Italia), 5-9 Septiembre 2005, pp. 39–43.

[112] S. L. Salecker, C. S. Hood, "Traffic Analysis of a Diverse Network that Overflows Dual Mode Phones", *IEEE Wireless Communications and Networking Conference*, 23-28 Septiembre 2000, vol. 2, pp. 824–828.

[113] S. Lincke-Salecker, C. S. Hood, "Call Placement Algorithm Evaluation for Integrated Networks that Overflow Speech or Data", *IEEE VTS 53rd Vehicular Technology Conference*, 6-9 Mayo 2001, vol. 2, pp. 936–940.

[114] S. Lincke-Salecker, "Performance and Service Issues in Selecting Adaptive Placement as a Load Distribution Technique", *IEEE 59th Vehicular Technology Conference*, 17-19 Mayo 2004, vol. 4, pp. 2057–2061.

[115] S. Lincke-Salecker, C. S. Hood, "Packet Data Capacity in an Integrated Network", *IEEE VTS 54th Vehicular Technology Conference*, 7-11 Octubre 2001, vol. 3, pp. 1870–1874.

[116] S. J. Lincke, "Wireless Load Sharing with Heterogeneous Services and Adaptive Placement", *Wireless Telecommunications Symposium*, 6-7 Abril 2005, pp. 79–84.

[117] S. Lincke-Salecker, "Load Shared Integrated Networks", *5th European Personal Mobile Communications Conference*, 22-25 Abril 2003, pp. 225–229.

[118] S. J. Lincke, "Vertical Handover Policies for Common Radio Resource Management", *International Journal of Communication Systems*, vol. 18, nº 6, pp. 527–543, 15 Marzo 2005, Wiley Interscience, DOI: 10.1002/dac.715. URL: http://doi.wiley.com/10.1002/dac.715.

[119] S. Lincke-Salecker, C. S. Hood, "A Supernet: Engineering Traffic Across Network Boundaries", *36th Annual Simulation Symposium*, 30 Marzo - 2 Abril 2003, pp. 117–124.

[120] A. Tölli, P. Hakalin, H. Holma, "Performance Evaluation of Common Radio Resource Management (CRRM)", *IEEE International Conference on Communications*, 28 Abril - 2 Mayo 2002, vol. 5, pp. 3429–3433.

[121] A. Tölli, P. Hakalin, "Adaptive Load Balancing Between Multiple Cell Layers", *IEEE 56th Vehicular Technology Conference*, 24-28 Septiembre 2002, vol. 3, pp. 1691–1695.

[122] O. Yilmaz, A. Furuskär, J. Pettersson, A. Simonsson, "Access Selection in WCDMA and WLAN Multi-Access Networks", *IEEE 61st Vehicular Technology Conference*, 30 Abril - 1 Mayo 2005, vol. 4, pp. 2220–2224.

[123] R. Veronesi, "Multiuser Scheduling with Multi Radio Access Selection", *2nd International Symposium on Wireless Communication Systems*, Siena (Italia), 5-9 Septiembre 2005, pp. 455–459.

[124] E. Adamopoulou, K. Demestichas, A. Koutsorodi, M. Theologou, "Intelligent Access Network Selection in Heterogeneous Networks", *2nd International Symposium on Wireless Communication Systems*, Siena (Italia), 5-9 Septiembre 2005, pp. 279–283.

[125] A. Pillekeit, F. Derakhshan, E. Jugl, A. Mitschele-Thiel, "Force-Based Load Balancing in Co-Located UMTS/GSM Networks", *IEEE 60th Vehicular Technology Conference*, 26-29 Septiembre 2004, vol. 6, pp. 4402–4406.

[126] R. Agustí, O. Salient, J. Pérez-Romero, L. Giupponi, "A Fuzzy-Neural Based Approach for Joint Radio Resource Management in a Beyond 3G Framework", *First International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks*, 2004, pp. 216–224.

[127] F. Malavasi, M. Breveglieri, L. Vignali, P. Leaves, J. Huschke, "Traffic Control Algorithms for a Multi Access Network Scenario Comprising GPRS and UMTS", *The 57th IEEE Semiannual Vehicular Technology Conference*, 22-25 Abril 2003, vol. 1, pp. 145–149.

[128] A. Furuskär, "Allocation of Multiple Services in Multi-Access Wireless Systems", *4th International Workshop on Mobile and Wireless Communications Network*, 9-11 Septiembre 2002, pp. 261–265.

[129] A. Furuskär, J. Zander, "Multiservice Allocation for Multiaccess Wireless Systems", *IEEE Transactions on Wireless Communications*, vol. 4, nº 1, pp. 174–184, Enero 2005.

[130] E. Mohyeldin, N. Motte, J. Luo, M. Dillinger, "Adaptive Radio Multihoming for Reconfigurable Terminals", *IST Mobile Communications Summit*, Aveiro (Portugal), Junio 2003.

[131] R. Mukerjee, J. Luo, M. Dillinger, E. Mohyeldin, "Joint Scheduling Algorithms over Heterogeneous Networks in a Reconfigurable Environment", *5th European Personal Mobile Communications Conference*, 22-25 Abril 2003, pp. 591–596.

[132] A. Y. T. Chung, M. Hassan, "Traffic Distribution Schemes for Multi-homed Mobile Hotspots", *IEEE 61st Vehicular Technology Conference*, 30 Abril - 1 Mayo 2005, vol. 4, pp. 2127–2131.

[133] D. K. Goldenberg, L. Qiuy, H. Xie, Y. R. Yang, Y. Zhang, "Optimizing Cost and Performance for Multihoming", *ACM SIGCOMM Computer Communication Review*, vol. 34, nº 4, pp. 79–92, Octubre 2004.

[134] EU IST FP5, Proyecto DRiVE (*Dynamic Radio for IP-Services in Vehicular Environments*), URL: http://www.ist-drive.org.

[135] EU IST FP5, Proyecto OverDRiVE (*Spectrum Efficient Uni- and Multicast Services over Dynamic Multi-Radio Networks in Vehicular Environments*), URL: http://www.ist-overdrive.org.

[136] Proyecto DRiVE, Deliverable D9, "Dynamic Spectrum Allocation Algorithm including Results of DSA Performance Simulations", URL: http://www.ist-drive.org/Deliverables/DRiVE_WP1D09pdf_v1.0_formatted.pdf.

[137] Proyecto OverDRiVE, Deliverable D13, "Specification and Performance of Dynamic Spectrum Allocation including Evaluation of Spectrum Coexistence", URL: http://www.comnets.rwth-aachen.de/~o_drive/publications/overdrive-wp1-D13-v1.0.pdf.

[138] Proyecto DRiVE, Deliverable D4.5, "Functional Description of Traffic Control in DRiVE", URL: http://www.ist-drive.org/Deliverables/DRiVE_WP2D04_pdf.zip.

[139] Proyecto DRiVE, Deliverable D12, "Description of Traffic Distribution Process Suitable for DRiVE", URL: http://www.ist-drive.org/Deliverables/drive_wp1d12_v100_pdf.zip.

[140] EU IST FP6, Proyecto EVEREST (*Evolutionary Strategies for Radio Resource Management in Cellular Heterogeneous Networks*), URL: http://www.everest-ist.upc.es.

[141] Proyecto EVEREST, Deliverable D11, "First Report on the Evaluation of RRM-CRRM Algorithms", URL: http://www.everest-ist.upc.es/registro.php?file=WP3-UPC-D11-Apr-001.pdf.

[142] Proyecto EVEREST, Deliverable D15, "Report on the Evaluation of RRM-CRRM Algorithms", URL: http://www.everest-ist.upc.es/registro.php?file=WP3-UPC-D15-Apr-001.pdf.

[143] EU IST FP6, Proyecto *Ambient Networks*, URL: http://www.ambient-networks.org.

[144] EU IST FP5, Proyecto MONASIDRE (*Management Of Networks And Services In a Diversified*), URL: http://www.monasidre.com.

[145] WWI, Proyecto WINNER (*Wireless World Initiative New Radio*), URL: https://www.ist-winner.org/.

[146] 3GPP, *Improvement of RRM across RNS and RNS/BSS (Release 5), 3GPP TR 25.881, versión 5.0.0*, 2001.

[147] 3GPP, *Improvement of RRM across RNS and RNS/BSS (Post Rel-5); (Release 6), 3GPP TR 25.891, versión 0.3.0*, 2003.

[148] 3GPP, *Feasibility study on 3GPP system to Wireless Local Area Network (WLAN) interworking (Release 6), 3GPP TR 22.934, versión 6.2.0*, 2003.

[149] 3GPP, *3GPP system to Wireless Local Area Network (WLAN) interworking; System description (Release 7), 3GPP TS 23.234, versión 7.0.0*, 2005.

[150] 3GPP, *3GPP system to Wireless Local Area Network (WLAN) interworking; Functional and architectural definition (Release 6), 3GPP TR 23.934, versión 1.0.0*, 2005.

[151] 3GPP, *Wireless Local Area Network (WLAN) interworking security (Release 6), 3GPP TS 33.234, versión 6.7.0*, 2005.

[152] M. López-Benítez, J. Gozálvez, "QoS Provisioning in Beyond 3G Heterogeneous Wireless Systems through Common Radio Resource Management Algorithms", *2nd ACM Workshop on QoS and Security for Wireless and Mobile Networks (Q2SWinet 2006)*, Torremolinos, 2-6 Octubre 2006, aceptado para su publicación.

[153] J. Laiho, A. Wacker, T. Novosad, *Radio Network Planning and Optimisation for UMTS*, John Wiley & Sons, 1999.

[154] 3GPP, *Services and service capabilities, 3GPP TS 22.105, versión 6.3.0*, 2005.

[155] M. López Benítez, J. Gozálvez, M. C. Lucas Estañ, "SPHERE - Plataforma Avanzada de Simulación para Sistemas Inalámbricos Heterogéneos", *XXI Simpósium Nacional de la Unión Científica Internacional de Radio*, Oviedo, 12-15 Septiembre 2006, aceptado para su publicación.

[156] M. López-Benítez, M. C. Lucas-Estañ, J. Gozálvez, "A Dynamic Radio Simulation Platform for the Study of Radio Resource Management Techniques in Heterogeneous Wireless Systems", *The 9-th ACM/IEEE International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, Torremolinos, 2-6 Octubre 2006, aceptado para su publicación.

[157] M. Junius, M. Steppler, M. Büter, D. Pesch, "Communication Networks Class Library (CNCL)", ComNets, Aachen University of Technology, 1996, URL: http://www.comnets.rwth-aachen.de.

[158] Y. Okumura, E. Ohmori, T. Kawano, K. Fukuda, "Field Strength and its Variability in the VHF and UHF Land Mobile Service", *Review Electronic Communication Lab.*, vol. 16, nº 9-10, pp. 825–873, 1968.

[159] M. Hata, "Empirical Formula for Propagation Loss in Land Mobile Radio Services", *IEEE Transactions on Vehicular Technology*, vol. VT-29, nº 3, pp. 317–325, Agosto 1980.

[160] COST 231, "Urban Transmission Loss Models for Mobile Radio in the 900 and 1800 MHz Bands", Septiembre 1991, TD(91)73.

[161] K. Feher, *Wireless Digital Communications: Modulation and Spread Spectrum Applications*, Prentice Hall, 1995.

[162] M. Gudmundson, "Correlation Model for Shadow Fading in Mobile Radio Systems", *IEEE Electronics Letters*, vol. 27, nº 23, pp. 2145–2146, Noviembre 1991.

[163] R. Gollreiter, "Channel Models Issue 2, Technical Report R2084-esg-cc3-ds-p-029-b1", , RACE 2084 Advanced TDMA Mobile Access, 1994.

[164] J. Pons-Puig, *Application of Link Adaptation to Evolved TDMA Mobile Communications Systems*, Tesis doctoral, University of Strathclyde, 1999.

[165] C. Passerini, G. Falciasecca, "Correlation between Delay-Spread and Orthogonality Factor in Urban Environments", *Electronics Letters*, vol. 37, nº 6, pp. 384–386, 15 Marzo 2001.

[166] H. Holma, A. Toskala, *WCDMA for UMTS*, Wiley, 2ª edición, 2002.

[167] T. Berners-Lee, R. Fielding, H. Frystyk, *Hypertext Transfer Protocol HTTP/1.0, RFC 1945, Network Working Group*, Mayo 1996.

[168] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, T. Berners-Lee, *Hypertext Transfer Protocol HTTP/1.1, RFC 2068, Network Working Group*, Enero 1997.

[169] P. Barford, M. Crovella, "Generating Representative Web Workloads for Network and Server Performance Evaluation", *Joint International Conference on Measurement and Modeling of Computer Systems*, Madison, Junio 1998, pp. 151–160.

[170] J. Ho, Y. Zhu, S. Madhavapeddy, "Throughput and Buffer Analysis for GSM General Packet Radio Service (GPRS)", *IEEE Wireless Communications and Networking Conference*, Nueva Orleans (Estados Unidos), Septiembre 1999, vol. 3, pp. 1427–1431.

[171] S. N. Fabri, S. Worrall, A. Kondoz, A. Sadka, "Video Communications over Mobile Networks", Communicate 2000, conferencia *on-line*, Octubre 2000, URL: http://www.ee.surrey.ac.uk/Personal/S.Worrall/Publications/comm2k.pdf.

[172] G. Côté, L. Winger, "Recent Advances in Video Compression Standards", *IEEE Canadian Review*, , nº 40, pp. 21–24, 2002.

[173] S. N. Fabri, S. Worrall, A. Sadka, A. Kondoz, "Real-Time Video Communications over GPRS", *IEE 3G Mobile Communication Technologies*, Mayo 2000, pp. 426–430.

[174] ITU-T, *Video Coding for Low Bit Rate Communication, Recomendación H.263*, 1998.

[175] K. Rijkse, "H.263: Video Coding for Low-Bit-Rate Communication", *IEEE Communications Magazine*, vol. 34, nº 12, pp. 42–45, Diciembre 1996.

[176] B. Girod, N. Färber, E. Steinbach, "Performance of the H.263 Video Compression Standard", *Journal on VLSI Signal Processing: System Signal, Image and Video Technology*, vol. 17, pp. 101–111, Noviembre 1997.

[177] G. Côté, B. Erol, M. Gallant, F. Kossentini, "H.263+: Video Coding at Low Bit Rates", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, nº 7, pp. 849–866, Noviembre 1998.

[178] K. Chawla, Z. Jiang, X. Qiu, A. Reibman, "Transmission of Streaming Video over an EGPRS Wireless Network", *IEEE International Conference on Multimedia*, Julio 2000, vol. 1, pp. 275–278.

[179] K. Stuhlmüller, N. Färber, M. Link, B. Girod, "Analysis of Video Transmission over Lossy Channels", *IEEE Journal on Selected Areas in Communications*, vol. 18, nº 6, pp. 1012–1032, Junio 2000.

[180] O. Lázaro, *Dynamic Radio Resource Management Algorithms and Traffic Models for Emerging Mobile Communication Systems*, Tesis doctoral, University of Strathclyde, 2001.

[181] O. Lázaro, D. Girma, J. Dunlop, "H.263 Video Traffic Modelling for Low Bit Rate Wireless Communications", *IEEE International Sysmposium on Personal, Indoor and Mobile Radio Communications*, Barcelona, Septiembre 2004, pp. 2124–2128.

[182] "MPEG-4 and H.263 Video Traces for Network Performance Evaluation", URL: http://www-tkn.ee.tu-berlin.de/research/trace/trace.html.

[183] ETSI-SMG, *EDGE Feasibility Study, Work item 184; Improved Data Rates through Optimized Modulation, Tdoc 97-331*, Diciembre 1997.

[184] M. López-Benítez, J. Gozálvez, O. Lázaro, "Algoritmos de Link Adaptation para la Transmisión de Datos en Tiempo Real en Redes de Comunicaciones Móviles", *XX Simpósium Nacional de la Unión Científica Internacional de Radio*, Gandía, 14-16 Septiembre 2005, pp. 99–100.

[185] J. Gozálvez, M. López-Benítez, O. Lázaro, J. Evangelio, "Design and Evaluation of a Link Adaptation Algorithm for Enhanced Video Service Provision within Mobile Radio Environments", *15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Barcelona, 5-8 Septiembre 2004, vol. 1, pp. 355–359.

[186] J. Gozálvez, M. López-Benítez, O. Lázaro, "Link Adaptation Algorithm for Improved Wireless Transmission of Delay-Sensitive Packet Data Services", *Electronics Letters*, vol. 41, nº 14, pp. 813–815, 7 Julio 2005.

[187] J. Gozálvez, *Performance Analysis of Link Adaptation for Circuit and Packet Switched TDMA Mobile Communications Systems*, Tesis doctoral, Department of Electronic and Electrical Engineering, University of Strathclyde, 2002.

[188] F. Frederiksen, T. E. Kolding, "Performance and Modeling of WCDMA/HSDPA Transmission/H-ARQ Schemes", *IEEE 56th Vehicular Technology Conference*, 24-28 Septiembre 2002, vol. 1, pp. 472–476.

[189] EU IST FP5, EURANE (*Enhanced UMTS Radio Access Network Extensions for ns-2*), "User Guide (Release 1.6)", URL: http://www.ti-wmc.nl/eurane.

[190] P. J. Ameigeiras Gutiérrez, J. Wigard, P. N. Andersen, H. C. Damgaard, P. Mogensen, "Performance of Link Adaptation in GPRS Networks", *IEEE Vehicular Technology Conference*, Septiembre 2000, vol. 2, pp. 492–499.

[191] B. Wang, K. I. Pedersen, T. E. Kolding, P. E. Mogensen, "Performance of VoIP on HSDPA", *IEEE 61st Vehicular Technology Conference*, 30 Mayo - 1 Junio 2005, vol. 4, pp. 2335–2339.

[192] M. Andersin, N. Mandayam, R. Yates, "Subspace Based Estimation of the Signal to Interference Ratio for TDMA Cellular Systems", *IEEE 46th Vehicular Technology Conference*, Atlanta (Estados Unidos), Abril 1996, pp. 1155–1159.

[193] M. Andersin, N. Mandayam, R. Yates, "Real-Time Estimation of the Signal to Interference Ratio in Cellular Radio Systems", *IEEE 47th Vehicular Technology Conference*, Phoenix (Estados Unidos), Mayo 1997, pp. 1089–1093.

[194] ETSI, *Digital Cellular Telecommunications System (Phase 2+); Physical Layer on the Radio Path; General Description, GSM 05.01, versión 8.4.0*, 1999.

[195] ETSI, *Digital Cellular Telecommunications System (Phase 2+); Multiplexing and Multiple Access on the Radio Path, GSM 05.02, versión 8.5.1*, 1999.

[196] ETSI, *Digital Cellular Telecommunications System (Phase 2+); Channel Coding, GSM 05.03, versión 8.5.1*, 1999.

[197] ETSI, *Digital Cellular Telecommunications System (Phase 2+); Modulation, GSM 05.04, versión 8.1.2*, 1999.

[198] ETSI, *Digital Cellular Telecommunications System (Phase 2+); General Packet Radio Service (GPRS); Overall Description of the GPRS Radio Interface; Stage 2, 3GPP TS 03.64, versión 8.10.0*, 1999.

[199] ETSI, *Digital Cellular Telecommunications System (Phase 2+); General Packet Radio Service (GPRS) Service Description; Stage 2, 3GPP TS 03.60, versión 7.9.0*, 1998.

[200] ETSI, *Digital Cellular Telecommunications System (Phase 2+); General Packet Radio Service (GPRS); Mobile Station - Base Station System (BSS) Interface; Radio Link Control/Medium Access Control (RLC/MAC) Protocol, 3GPP TS 04.60, versión 8.17.0*, 1999.

[201] S. Parkvall, E. Dahlman, P. Frenger, P. Beming, M. Persson, "The Evolution of WCDMA towards Higher Speed Downlink Packet Data Access", *IEEE VTS 53rd Vehicular Technology Conference*, 6-9 Mayo 2001, vol. 3, pp. 2287–2291.

[202] T. E. Kolding, F. Frederiksen, P. E. Mogensen, "Performance Aspects of WCDMA Systems with High Speed Downlink Packet Access (HSDPA)", *IEEE 56th Vehicular Technology Conference*, 24-28 Septiembre 2002, vol. 1, pp. 477–481.

[203] J. H. Horng, G. Vannucci, J. Zhang, "Capacity Enhancement for HSDPA in WCDMA System", *IEEE 56th Vehicular Technology Conference*, 24-28 Septiembre 2002, vol. 2, pp. 661–665.

[204] R. Love, A. Ghosh, W. Xiao, R. Ratasuk, "Performance of 3GPP High Speed Downlink Packet Access (HSDPA)", *IEEE 60th Vehicular Technology Conference*, 26-29 Septiembre 2004, vol. 5, pp. 3359–3363.

[205] C.-S. Chiu, C.-C. Lin, "Comparative Downlink Shared Channel Evaluation of WCDMA Release 99 and HSDPA", *IEEE International Conference on Networking, Sensing and Control*, 2004, vol. 2, pp. 1165–1170.

[206] 3GPP, *UTRA High Speed Downlink Packet Access, 3GPP TR 25.950, versión 4.0.1*, 2005.

[207] 3GPP, *Physical Layer Aspects of UTRA High Speed Downlink Packet Access, 3GPP TR 25.848, versión 4.0.0*, 2001.

[208] 3GPP, *Physical Channels and Mapping of Transport Channels onto Physical Channels (FDD), 3GPP TS 25.211, versión 6.7.0*, 2005.

[209] 3GPP, *Multiplexing and Channel Coding (FDD), 3GPP TS 25.212, versión 6.7.0*, 2005.

[210] 3GPP, *Spreading and Modulation (FDD), 3GPP TS 25.213, versión 6.4.0*, 2005.

[211] 3GPP, *Physical Layer Procedures (FDD), 3GPP TS 25.214, versión 6.7.1*, 2005.

[212] 3GPP, *Physical Layer - Measurements (FDD), 3GPP TS 25.215, versión 6.4.0*, 2005.

[213] 3GPP, *Physical Channels and Mapping of Transport Channels onto Physical Channels (TDD), 3GPP TS 25.221, versión 6.5.0*, 2005.

[214] 3GPP, *Multiplexing and Channel Coding (TDD), 3GPP TS 25.222, versión 6.2.0*, 2005.

[215] 3GPP, *Spreading and Modulation (TDD), 3GPP TS 25.223, versión 6.1.0*, 2005.

[216] 3GPP, *Physical Layer Procedures (TDD), 3GPP TS 25.224, versión 6.6.0*, 2005.

[217] 3GPP, *Physical Layer - Measurements (TDD), 3GPP TS 25.225, versión 6.1.0*, 2005.

[218] 3GPP, *High Speed Downlink Packet Access: Physical Layer Aspects, 3GPP TR 25.858, versión 5.0.0*, 2002.

[219] T. E. Kolding, F. Frederiksen, P. E. Mogensen, "Performance Evaluation of Modulation and Coding Schemes Proposed for HSDPA in 3.5G UMTS Networks", *International Symposium on Wireless Personal Multimedia Communications*, Aalborg (Dinamarca), September 2001, vol. 1, pp. 307–312.

[220] 3GPP, *Report on Hybrid ARQ Type II/III, 3GPP TR 25.835, versión 1.0.0*, 2000.

[221] D. Chase, "Code Combining - A Maximum Likelihood Decoding Approach for Combining an Arbitrary Number of Noise Packets", *IEEE Transactions on Communications*, vol. 33, pp. 441–448, Mayo 1985.

[222] 3GPP, *Radio Interface Protocol Architecture, 3GPP TS 25.301, versión 6.4.0*, 2005.

[223] 3GPP, *Services Provided by the Physical Layer, 3GPP TS 25.302, versión 6.5.0*, 2005.

ÍNDICE ALFABÉTICO

A

Adaptive Modulation and Coding (AMC) 92, 160
Advanced Mobile Phone Service (AMPS) 25, 26
Algoritmo
  UBET (Utility Based on Expected Throughput) 53, 54
  UBReQoS (Utility Based on Required Quality of Service) . . 53, 62
Aproximación para calcular CIR/SIR 58
Asignación dinámica de espectro 31
Automatic Repeat reQuest (ARQ) 152

B

Backward Error Correction (BEC) 165
BLock Error Rate (BLER) 54, 76
Block Sequence Number (BSN) 154
Bluetooth 26, 36

C

Código
  de aleatorización 161
  de canalización 161
Carrier-to-Interference Ratio (CIR) 55, 59, 76, 150
CDMA
  1xRTT 26
cdma2000 26
  1X-EV-DV/DO 26
cdmaOne 26
Channelisation code 161
Chase combining 167
Chip 160
Code Division Multiple Access (CDMA) . . . 26
Coding Scheme (CS) 149
Common Radio Resource Management (CRRM) 31
Communication Networks Class Library (CNCL) 77
Corrección de errores
  hacia atrás 165
  hacia delante 165
Correo electrónico
  modelo de tráfico 69, 85
Cumulative Distribution Function (CDF) 69, 91

D

Digital Audio Broadcasting (DAB) 26
Digital Broadcasting System (DBS) 26
Digital Enhanced Cordless Telecommunications (DECT) 26
Digital Video Broadcasting (DVB) 26
Direct-Sequence Code Division Multiple Access (DS-CDMA) 160
Downlink 148
Downlink Shared CHannel (DSCH) 160
Dynamic Spectrum Allocation (DSA) 31

E

Efecto ping-pong 43
Enhanced Data-rates for GSM/Global Evolution (EDGE) 25, 26, 147
Enlace
  ascendente 148
  descendente 148
European Telecommunications Standardization Institute (ETSI) 148

F

Factor de ortogonalidad 61, 83
Fast Cell Selection (FCS) 160
Fast scheduling 160
First-Come First-Served (FCFS) 39, 81
First-In First-Out (FIFO) 39
Forward Error Correction (FEC) 165
Frequency Division Duplex (FDD) 148
Frequency Division Multiple Access (FDMA) 148

Función de utilidad
U_1
algoritmo UBET 55, 56
algoritmo UBReQoS 65
U_2
algoritmo UBET 57, 58
algoritmo UBReQoS 67
U_3
algoritmo UBReQoS 67
u_1
algoritmo UBReQoS 63, 64
esperada de
von Neumann–Morgenstern 65

G

Gaussian Minimum Shift Keying (GMSK) 149
General Packet Radio Service (GPRS) 25, 26, 147
Global Positioning System (GPS) 30
Global System for Mobile communications (GSM) 25, 26
GSM/EDGE Radio Access Network (GERAN) 36

H

Handover
horizontal 32
vertical 32
HIgh PErformance Radio Local Area Network Type 2 (HIPERLAN/2) 26
HIgh PErformance Radio Metropolitan Area Network (HIPERMAN) 26
High Speed Circuit Switched Data (HSCSD) 25, 26
High Speed Downlink Packet Access (HSDPA) 147, 160
Hot spot 30, 32
Hybrid Automatic Repeat reQuest (HARQ) 160
tipo I 166
tipo II (Full IR) 166
tipo III 166
Chase combining 167
Partial IR 167
HyperText Transfer Protocol (HTTP) 85

I

Incremental Redundancy (IR) 166
Full (Full IR) 166
Partial (Partial IR) 167
Information Society Technologies (IST) 49
Interferencia
intercelda 61, 83
intracelda 61, 83
International Standardization Organization (ISO) 88
International Telecommunication Union (ITU) 88
IS-136 26
IS-54 26
IS-95 26

L

Link Adaptation (LA) 92
Look-Up Table (LUT) 76

M

MAGIC, concepto 27
Medium Access Control (MAC) 81, 168
Mobile Broadband System (MBS) 50
Modelo
de atenuación (pathloss) 79
de calidad del canal
en GPRS y EDGE 95
en HSDPA 95
de efecto sombra (shadowing) 81
de movilidad 82
de tráfico 84
de correo electrónico 69, 85
de navegación web 69, 85
de vídeo H.263 88
Modo de transmisión 63, 65
Multiple Input Multiple Output (MIMO) . 160

N

Navegación web
modelo de tráfico 69, 85
Nivel
de enlace 76
de sistema 76
Nordic Mobile Telephone (NMT) 25, 26

O

Okumura-Hata
modelo de propagación 79
extensión COST 231 . . . . . 58, 79
Open System Interconnection (OSI) 152

P

Personal Digital Cellular (PDC) 26
Personal Handyphone System (PHS) 26
Plataforma de simulación SPHERE 77

Q

Quality of Service (QoS) 54

R

Radio Access (RA) 33
Radio Access Network (RAN) 33
Radio Access Technology (RAT) 33
Radio Link Control (RLC) 152
Radio multihoming 47

Red
heterogénea . . . 29
inalámbrica
de área local (WLAN) . . . 26
de área metropolitana (WMAN) 26
de área personal (WPAN) . . . 26
Redundancia incremental . . . 166
parcial . . . 167
total . . . 166
Relación
portadora a interferencia
(CIR) . . . 55, 59, 76, 150
señal a interferencia
(SIR) . . . 55, 61, 76, 163
señal a ruido (SNR) . . . 41
Rendimiento
algoritmo UBET . . . 106
algoritmo UBReQoS . . . 115

S

Scheduling . . . 168
Maximum CIR . . . 169
Round Robin . . . 168
Scrambling code . . . 161
Secuencia
de aleatorización . . . 161
de canalización . . . 161
Sequence Number Space (SNS) . . . 154
Signal-to-Interference Ratio
(SIR) . . . 55, 61, 76, 163
Signal-to-Noise Ratio (SNR) . . . 41
Simulación
a nivel de enlace . . . 76
a nivel de sistema . . . 76
SPHERE, plataforma de simulación . . . 77
Stop-and-Wait (SAW) . . . 169

T

Terminal de usuario
multi-estándar . . . 29
reconfigurable . . . 29
Time Division - Synchronous Code Division
Multiple Access (TD-SCDMA) . . . 26
Time Division Multiple Access
(TDMA) . . . 26, 148
Total Access Communication System
(TACS) . . . 25, 26
Trama de vídeo H.263
tipo I . . . 88
tipo P . . . 88
tipo PB . . . 89
Transmission Control Protocol (TCP) . . . 85
Transmission Mode (TM) . . . 63, 65
Transmission Time Interval (TTI) . . . 161
Transport Block (TB) . . . 168
Transport Block Set (TBS) . . . 168
Traspaso
horizontal . . . 32
vertical . . . 32

U

UBET
algoritmo (Utility Based on
Expected Throughput) . . . 53, 54
UBReQoS
Algoritmo (Utility Based on
Required Quality of Service) . . 53, 62
UMTS Terrestrial Radio Access Network
(UTRAN) . . . 36
Universal Mobile Telecommunications System
(UMTS) . . . 25, 26
Uplink . . . 148

V

Variación de la demanda
espacial . . . 30
temporal . . . 30
Video H.263
modelo de tráfico . . . 88
trama de
tipo I . . . 88
tipo P . . . 88
tipo PB . . . 89

W

Web
modelo de tráfico de navegación . . . 85
Wideband Code Division Multiple Access
(WCDMA) . . . 26
Window Size (WS) . . . 154
Wireless Local Area Network (WLAN) . . . 26
Wireless Metropolitan Area Network
(WMAN) . . . 26
Wireless Personal Area Network (WPAN) . . 26
Worldwide Interoperability for Microwave Access (WiMAX) . . . 26
Wrap-arround . . . 77