

LTE



LTE: NUEVAS TENDENCIAS EN COMUNICACIONES MÓVILES

Coordinador

Ramon Agusti Comes

Autores

Ramon Agusti Comes
Francisco Bernardo Álvarez
Fernando Casadevall Palacio
Ramon Ferrús Ferre
Jordi Pérez Romero
Oriol Sallent Roig



LTE:
NUEVAS TENDENCIAS
EN COMUNICACIONES
MÓVILES

AUTORES

Ramón Agustí, Francisco Bernardo, Fernando Casadevall,
Ramon Ferrús, Jordi Pérez-Romero, Oriol Sallent

Coordinador

Ramón Agusti Comes

Catedrático de Universidad
Departamento de Teoría de la Señal y Comunicaciones
Universitat Politècnica de Catalunya

Autores

Ramón Agusti Comes

Catedrático de Universidad
Departamento de Teoría de la Señal y Comunicaciones
Universitat Politècnica de Catalunya

Francisco Bernardo Álvarez

Ingeniero de Telecomunicación
Departamento de Teoría de la Señal y Comunicaciones
Universitat Politècnica de Catalunya

Fernando Casadevall Palacio

Catedrático de Universidad
Departamento de Teoría de la Señal y Comunicaciones
Universitat Politècnica de Catalunya

Ramón Ferrús Ferré

Profesor Titular de Universidad
Departamento de Teoría de la Señal y Comunicaciones
Universitat Politècnica de Catalunya

Jordi Pérez Romero

Profesor Titular de Universidad
Departamento de Teoría de la Señal y Comunicaciones
Universitat Politècnica de Catalunya

Oriol Sallent Roig

Profesor Titular de Universidad
Departamento de Teoría de la Señal y Comunicaciones
Universitat Politècnica de Catalunya

Edita: Fundación Vodafone España

ISBN: 84-934740-4-5

D.L.: M-34503-2010

© Copyright 2010. Fundación Vodafone España.

Queda prohibida cualquier reproducción total o parcial de este libro y del CD-Rom, almacenamiento en un sistema informático o transmisión en cualquier forma o por cualquier medio (electrónico, fotocopia y otros métodos).

PRÓLOGO

La Fundación Vodafone España siempre ha tenido una especial preocupación por potenciar el conocimiento sobre los fundamentos tecnológicos de la telefonía móvil en nuestras universidades y el sector especializado que desee ponerse al día con rigor, en una tecnología compleja que ha devenido, mediante un progreso tecnocientífico continuado y de intensidad creciente, en utensilio imprescindible para el trabajo y las relaciones que desarrollan la mayor parte de los ciudadanos. El objetivo ha sido claro, la tecnología que en la actualidad soporta la comunicación más utilizada hoy, no es baladí. Se fundamenta en un sustrato teórico que, a gran velocidad, va ofreciendo desde la investigación, el desarrollo y la innovación, soluciones crecientemente integradoras, crecientemente flexibles y versátiles.

En 1999 la entonces Fundación Airtel Móvil, hoy Fundación Vodafone España, comenzó una serie de publicaciones en el sentido indicado que han continuado hasta el presente. En aquel año, presentamos una obra pionera y básica, "Primeros Pasos de la Telecomunicación", del Profesor Sebastián Olivé que recogía las lecciones y conferencias que desde la Cátedra Airtel se impartieron en el curso 1997-1998. En ella se ofrecía una completa descripción del proceso de avance de las telecomunicaciones en España.

Ese mismo año de 1999 publicábamos "Comunicaciones Móviles GSM", bajo la coordinación del Profesor José María Hernando Rábanos, Catedrático de la Universidad Politécnica de Madrid, primera obra en español que ofrecía una completa descripción del sistema GSM así como de sus avances desde su lanzamiento. Esa obra fue fundamental y de referencia en el mundo por entonces todavía emergente de las comunicaciones móviles. Ofrecía junto a una elevada claridad y extensión didáctica, una potente herramienta de inmersión en una tecnología que surgía con fuerza en el panorama técnico, económico y social español. Colaboraron en ella un conjunto de jóvenes profesores y tecnólogos.

En el año 2002, el profesor Miguel Calvo Ramón, así mismo Catedrático de la UPM, presentaba la obra que llevaba por título "Sistemas de Comunicaciones Móviles de Tercera Generación IMT-2000 (UMTS)" editada por la Fundación Vodafone España, que contó con un numeroso grupo de colaboradores, varios de ellos ingenieros de esta Fundación.

En el año 2004 vio la luz una nueva publicación. Esta vez coordinada por el Profesor Ramón Agustí, Catedrático de la Universidad Politécnica de Cataluña, Director del Master de Comunicaciones Móviles que se ha venido impartiendo en Barcelona durante catorce años, y muy próximo a nuestra Fundación. Ya por entonces los servicios basados en los sistemas de comunicaciones 3G eran una realidad. Ramón Agustí y sus colaboradores nos propusieron ordenar el conocimiento sobre el tema, sistematizarlo y transformarlo, mediante un libro, en un elemento referencial y didáctico de primer nivel. Así se planificó esta obra que recopilaba las últimas actualizaciones del sistema UMTS y aparecieron, con la colaboración de los profesores Álvarez-Campana y Sallent, los “Elementos de arquitectura y gestión de recursos en redes UMTS”, un hito pedagógico y de consulta para los profesionales interesados en profundizar sus conocimientos en temas relativos al diseño y arquitectura de estas redes.

Nuevamente, nuestro entrañable amigo el Profesor Agustí hace un valioso regalo a la comunidad científica y tecnológica: un libro de anticipación sobre el futuro de la telefonía móvil. Según nos anuncia el autor en su prólogo, estamos en los albores de una nueva generación de comunicaciones móviles, la cuarta generación o 4G, de la que el sistema LTE (“Long Term Evolution”) puede considerarse el primer eslabón en una supuesta trayectoria evolutiva que debe culminar con el LTE-Advanced.

Ciertamente, la industria, los operadores de telefonía y los usuarios aguardan, con interés creciente, que se puedan eliminar definitivamente las barreras que todavía impedían la consecución plena de una movilidad con capacidad multimedia. Y en la implementación de la LTE reside la mayor esperanza de que ello sea posible.

Esto equivaldría a afirmar en palabras del propio Agustí “que con la aparición de LTE los usuarios que lo deseen ya no tendrán que verse penalizados en su capacidad de comunicación por el hecho de ser móviles en lugar de fijos”. Y esta cuestión, por muchas razones que a los lectores de este libro no pueden serles indiferentes, es un asunto importante.

Escribir sobre este nuevo libro del Profesor Agustí, nos ha dado la posibilidad de narrar la aparición de un conjunto de publicaciones rigurosas sobre comunicación móvil, hecho que ha constituido y constituye uno de sus más preciados signos de identidad tecnocientífica de nuestra Fundación. Hablar en este país de Agustí, es en muchos aspectos relatar, tecnológicamente y sin concesiones a lo fácil, la evolución de la telefonía móvil en España.

En el libro que hoy nos honramos en prologar desde la Fundación Vodafone, “LTE: nuevas tendencias en comunicaciones móviles”, Ramón Agustí ha coordinado un selecto grupo de autores entre los que figura él mismo junto a Francisco Bernardo, Fernando Casadevall, Ramon Ferrús, Jordi Pérez-Romero y Oriol Sallent.

Como Ramón dice en su prólogo introductorio y sencillo: “Nuestra pretensión ha sido aunar por una parte el rigor expositivo que un sistema estandarizado como LTE requiere con la pedagogía necesaria para una correcta comprensión de sus principios básicos, pedagogía de la que desgraciadamente carecen las especificaciones técnicas que emanan de los estándares”. Y creemos que lo han logrado con creces.

A todos los autores, al Profesor José M^a Hernando (quien ha realizado la revisión final del proyecto) y en especial al Profesor Agustí, recién nombrado académico de la Real Academia de Ingeniería, muchas gracias y nuestra más sincera enhorabuena por el magnífico trabajo realizado, primero que edita desde esta nueva y merecida silla.

Y a las personas de la Fundación Vodafone España, que esta publicación y en otras anteriores han trabajado intensamente para hacerlas realidad, nuestro expreso reconocimiento por su valioso trabajo.

Francisco Román
Presidente
VODAFONE España

José Luis Ripoll
Presidente
Fundación Vodafone España

INTRODUCCIÓN

Después de ya casi dos décadas de prácticamente ininterrumpido crecimiento de las comunicaciones móviles, primero de la mano de GSM y últimamente con el despliegue definitivo de UMTS, estamos en los albores de una nueva generación de comunicaciones móviles, la cuarta generación o 4G, de la que el sistema LTE (“Long Term Evolution”), cuya primera especificación fue concluida por 3GPP a finales de 2008, puede considerarse el primer eslabón en una supuesta trayectoria evolutiva que debe culminar con el LTE-Advanced.

Son muchas las variantes que LTE introduce en relación a sistemas de comunicaciones móviles previos y que tendremos ocasión de tratar en este libro, sin embargo dos aspectos relevantes que cabría destacar son que en LTE, por primera vez, todos los servicios, incluida la voz, se soportan sobre el protocolo IP (Internet Protocol), y que las velocidades de pico de la interfaz radio se sitúan dentro del rango de 100 Mb/s y 1 Gb/s, ampliamente superiores a las conseguidas en los sistemas predecesores. Se espera que con LTE se puedan romper finalmente y definitivamente las barreras que todavía impedían la consecución plena de una movilidad con capacidad multimedia. Lo que sería equivalente a afirmar que con la aparición de LTE los usuarios que lo deseen ya no tendrán que verse penalizados en su capacidad de comunicación por el hecho de ser móviles en lugar de fijos.

A la vista de lo anterior, este libro pretende sentar las bases de los principios en los que se sustenta el sistema LTE. Nuestra pretensión ha sido aunar por una parte el rigor expositivo que un sistema estandarizado como LTE requiere con la pedagogía necesaria para una correcta comprensión de sus principios básicos, pedagogía de la que desgraciadamente carecen las especificaciones técnicas que emanan de los estándares. Con esta finalidad hemos estructurado el libro en los siete capítulos que se introducen a continuación.

En el Capítulo 1: *Acceso radio móvil: visión perspectiva y prospectiva*, se despliega, a modo de capítulo introductorio, el contexto dentro del que el sistema LTE debe inmiscuirse. Aspectos tecnológicos de base se alternan con otros correspondientes al entorno regulatorio, del que la gestión del espectro es su principal exponente, o de estandarización dentro del que se ha gestado el sistema. Se pretende que en definitiva al sistema LTE se le vea un acomodo natural dentro del ecosistema tecnológico en el que deberá desenvolverse, tanto desde un punto de perspectiva del mismo como de prospectiva hacia nuevos paradigmas que están emergiendo, y que el sistema LTE va a acabar usando en mayor o menor medida.

En el Capítulo 2: *Arquitectura del sistema LTE*, se proporciona una visión global de la estructura de un sistema LTE mediante una aproximación que arranca con descripciones genéricas de las arquitecturas de los sistemas celulares, en especial de los sistemas desarrollados en el 3GPP. Esta aproximación permite una mejor comprensión del sistema LTE al hacer explícitos sus orígenes, a la vez que proporciona una diferenciación clara del ámbito que cubre LTE en relación con los sistemas precedentes. Sobre esta base, se van desmenuzando posteriormente los elementos constituyentes tanto de la red de acceso como del núcleo de red del sistema LTE, haciendo especial hincapié en las funciones que abarcan cada uno de los nodos de la red, sus interfaces y los principales protocolos sobre los que se sustentan dichas interfaces.

En el Capítulo 3: *Marco de gestión de sesiones, movilidad y seguridad en LTE*, una vez introducidos los elementos de la arquitectura de LTE en el Capítulo anterior, se describe su operativa en cuanto a la gestión de sesiones, movilidad y seguridad, todas ellas funcionalidades básicas de cualquier sistema celular. Se procede entonces a una descripción de los conceptos inherentes a cada una de las citadas funcionalidades, incluyendo aspectos tales como el modelo de provisión de servicios de conectividad IP con control de QoS que ofrece LTE en lo que respecta a la gestión de sesiones, los principios de funcionamiento de la localización y el handover en lo que respecta a la gestión de movilidad y los servicios de control de acceso y confidencialidad soportados por LTE en lo que respecta a seguridad. Además de la descripción del planteamiento adoptado en LTE para sustentar dichas funciones, se incluye también una descripción de los principales procedimientos o mecanismos finales de ejecución.

En el Capítulo 4: *Tecnologías de nivel físico*, se desarrollan de modo autocontenido los principios básicos de las tecnologías de transmisión OFDM y MIMO contempladas en LTE. Se introducen primero los fundamentos y motivaciones del esquema de modulación OFDM, en tanto que presente en las técnicas de acceso múltiple OFDMA y SC-FDMA, usadas en LTE en el enlace descendente y ascendente respectivamente, y que se describen a continuación. Se hace especial hincapié también en las estrategias multi-antena MIMO que, si bien no son específicas de LTE, en este sistema han encontrado terreno abonado para su explotación. Se termina el Capítulo con la descripción de las estructuras conjuntas de transmisión y recepción OFDMA con múltiples antenas.

En el Capítulo 5: *Interfaz radio del sistema LTE*, se describen los ingredientes básicos de la interfaz aire contemplados en la primera especificación del sistema. Tomando como punto de partida la torre de protocolos de la interfaz aire, en lo que respecta al plano de control y plano de usuario, se van desmenuzando las distintas capas, funciones, servicios portadores, canales lógicos, canales de transporte y canales físicos con las consiguientes estructuras de tramas y señales de sincronización. Todo lo anterior aplica tanto al enlace ascendente como descendente. El capítulo finaliza con la descripción de los principales procedimientos propios de la operativa de la interfaz radio.

En el Capítulo 6: *Gestión de recursos radio y del espectro radioeléctrico*, se describen el conjunto de funciones y mecanismos encaminados a permitir un máximo aprovechamiento

de los recursos espectrales y de potencia disponibles en LTE, limitados por naturaleza, de manera que se garantice un determinado valor de los parámetros de calidad o QoS establecidos. En este contexto cobra especial relevancia la apropiada gestión del control de admisión y congestión, de las interferencias intercelulares, del "scheduling" de paquetes, de la movilidad, del balanceo de carga, del control de potencia y la transmisión discontinua. También se detalla a lo largo del Capítulo propuestas de algoritmos de gestión de recursos radio, ya que, aún no siendo objeto de estandarización, tienen una especial relevancia en el correcto funcionamiento del sistema. Por último, se describe la capacidad inherente que ofrece el sistema LTE para soportar una gestión flexible del espectro asignado.

En el Capítulo 7: *Redes heterogéneas*, se aborda un escenario real donde los sistemas LTE se deben interconectar con otra multiplicidad de sistemas, algunos de ellos estandarizados también por el 3GPP como sucede con GSM y UMTS, pero no necesariamente en otros casos como las redes CDMA2000, WiMAX o WLAN. Para ello se introducen primero unos fundamentos de "interworking" entre redes que nos permiten abstraernos de las redes concretas a interconectar. En este sentido se destacan primero los aspectos de escenarios y arquitecturas genéricas, terminales multimodo, los distintos niveles de interconexión, handover y mecanismos de optimización entre distintos sistemas, para posteriormente detallar ya de manera específica la interconexión de redes LTE con UMTS y GSM así como con redes no 3GPP.

Los Autores

CAPITULO 1

1 Acceso radio móvil: visión perspectiva y prospectiva	21
1.1 Introducción.....	21
1.2 Principios de evolución y expansión	23
1.2.1 Despliegue	25
1.2.2 Espectro radioeléctrico	32
1.2.3 Tecnología y eficiencia espectral.....	39
1.3 Estandarización.....	46
1.3.1 LTE.....	46
1.3.2 LTE Advanced	48
1.4 Referencias bibliográficas.....	50

CAPITULO 2

2 Arquitectura del sistema LTE	55
2.1 Introducción	55
2.2 Arquitectura genérica de los sistemas celulares	56
2.3 Arquitectura general de los sistemas 3GPP	58
2.4 Arquitectura del sistema LTE.....	60
2.5 Red de acceso evolucionada: E-UTRAN	62
2.5.1 Arquitectura de E-UTRAN.....	62
2.5.2 Entidades de red e interfaces	64
2.5.3 Protocolos.....	70
2.5.4 Comparativa E-UTRAN y UTRAN	76
2.6 Red troncal de paquetes evolucionada: EPC	78
2.6.1 Arquitectura de EPC.....	78
2.6.2 Entidades de red e interfaces	82
2.6.3 Protocolos.....	89

2.6.4 Configuraciones de la red EPC	98
2.6.5 Soporte de itinerancia (roaming)	100
2.7 IP Multimedia Subsystem (IMS).....	103
2.8 Equipos de usuario.....	107
2.9 Referencias bibliográficas.....	109

CAPITULO 3

3 Marco de gestión de sesiones, movilidad y seguridad en LTE.....	115
3.1 Introducción.....	115
3.2 Gestión de sesiones	115
3.2.1 Servicio de conectividad IP: conexiones PDN	115
3.2.2 Servicio portador EPS	118
3.2.3 Modelo de QoS.....	121
3.2.4 Control del servicio de conectividad: sistema PCC.....	126
3.2.5 Procedimientos de gestión de sesiones	132
3.3 Gestión de movilidad.....	146
3.3.1 Marco de gestión de movilidad.....	147
3.3.2 Gestión de la localización	150
3.3.3 Mecanismo de handover	151
3.3.4 Procedimientos de gestión de movilidad	153
3.4 Gestión de seguridad.....	161
3.4.1 Marco general de seguridad.....	162
3.4.2 Seguridad de acceso a la red.....	165
3.4.3 Seguridad en la infraestructura de red.....	168
3.4.4 Procedimientos de gestión de seguridad.....	171
3.5 Referencias bibliográficas.....	173

CAPITULO 4

4 Tecnologías de nivel físico	177
4.1 Fundamentos de OFDM	177

4.1.1 Esquemas de modulación y demodulación.....	179
4.1.2 Efecto de la dispersión del canal: concepto de prefijo cíclico	183
4.2 OFDMA como técnica de acceso múltiple para el enlace descendente	190
4.2.1 Procedimientos de RRM asociados a OFDMA.....	193
4.2.2 Sistemas celulares OFDMA.....	197
4.2.3 Parámetros de OFDMA empleados por LTE.....	200
4.3 SC-FDMA como técnica de acceso múltiple para el enlace ascendente	202
4.3.1 Motivación	202
4.3.2 Esquemas de transmisión y recepción SC-FDMA	204
4.3.3 Procedimientos de RRM asociados a SC-FDMA.....	210
4.3.4 Parámetros de SC-FDMA empleados por LTE.....	210
4.4 Estructuras con varias antenas.....	211
4.4.1 Caracterización de las estructuras MIMO	215
4.4.2 Formulación de las Estructuras MIMO	215
4.4.3 Capacidad de las Estructuras MIMO.....	219
4.4.4 Precodificación	222
4.4.5 MIMO multiusuario (MU MIMO).....	227
4.4.6 Comparativa entre estructuras MIMO y MU MIMO.....	231
4.5 Estructuras de transmisión y recepción OFDMA con múltiples antenas.....	231
4.6 Referencias bibliográficas.....	235
Anexo 4.1. Notación de señales complejas.....	237

CAPITULO 5

5 Interfaz Radio del sistema LTE	241
5.1 División funcional y Pila de protocolos	242
5.2 Radio Resource Control (RRC)	244
5.3 Packet Dependence Convergence Protocol (PDCP): Funcionalidades	
Principales.....	246
5.4 Capa RLC-MAC	246
5.4.1 Funcionalidades relevantes de la sub-capa RLC.....	246

5.4.2 Funcionalidades relevantes de la sub-capa MAC.....	248
5.5 Capa Física.....	251
5.5.1 Concepto de bloque de recursos físicos (Physical Resource Block).....	252
5.5.2 Estructura de trama	254
5.5.3 Enlace Descendente	256
5.5.4 Enlace Ascendente	273
5.6 Mapeo entre canales Físicos, de Transporte y Lógicos.....	288
5.7 Categorías de terminales móviles en LTE.....	290
5.8 Procedimientos básicos asociados a la interfaz aire.....	292
5.8.1 Procedimiento de sincronización y adquisición inicial de parámetros .	292
5.8.2 Procedimiento de acceso aleatorio	294
5.8.3 Mecanismo de aviso (Paging)	297
5.8.4 Mecanismo de establecimiento del servicio portador radio	298
5.8.5 Mecanismo de Información del estado del canal (CQI-Reporting).....	301
5.9 Referencias bibliográficas.....	303
Anexo 5.1. Secuencias utilizadas en la interfaz aire LTE.....	305
I.1 Secuencias de Zadoff–Chu	305
I.2 Códigos de Walsh-Hadamard	307
I.3 Secuencias de máxima longitud (m-sequences).....	309
I4.- Códigos de Gold	311
Anexo 5.2.- Sistema de Información en la interfaz aire LTE.....	313
Anexo 5.3.- Mecanismos de Retransmisión Híbrida (HARQ mechanisms).....	315
III.1 Redundancy Version (RV).....	317
Anexo 5.4. Servicio Multicast and Broadcast (MBMS)	319
Anexo 5,5. Mecanismos de seguridad en la interfaz aire.....	322

CAPITULO 6

6 Gestión de Recursos Radio y del Espectro Radioeléctrico	327
6.1 Introducción	327
6.2 Control de admisión radio.....	330

6.3 Control de servicios portadores radio	331
6.4 Scheduling de paquetes	332
6.4.1 Información de estado del canal.....	335
6.4.2 Información de estado del buffer	336
6.4.3 Información de disponibilidad de potencia.....	337
6.4.4 Adaptación dinámica del enlace.....	338
6.4.5 Estrategias de scheduling.....	340
6.4.6 Scheduling semi-persistente.....	343
6.5 Inter-Cell Interference Coordination	343
6.5.1 Indicadores de interferencia y estrategias ICIC en downlink.....	346
6.5.2 Indicadores de interferencia y estrategias ICIC en uplink.....	348
6.5.3 Otras técnicas de mitigación de interferencia intercelular	348
6.6 Control de la movilidad.....	348
6.6.1 Etapas para la ejecución de estrategias RRM de movilidad.	350
6.6.2 Estrategias de reelección de celda.....	352
6.6.3 Estrategias de handover	354
6.7 Balanceo de carga.....	357
6.8 Control de potencia.....	358
6.8.1 Control de potencia para el PUCCH	359
6.8.2 Control de potencia para el PUSCH	360
6.9 Recepción discontinua	361
6.10 Gestión flexible del espectro	364
6.10.1 Flexibilidad en la banda de despliegue	364
6.10.2 Flexibilidad en el ancho de banda en uso.....	365
6.10.3 Asignación dinámica de espectro	366
6.11 Referencias bibliográficas.....	367

CAPITULO 7

7 Redes Heterogéneas.....	373
7.1 Generalidades.....	373

7.2 Fundamentos de las soluciones de interworking	374
7.2.1 Escenario de referencia	374
7.2.2 Arquitectura genérica de las redes de acceso	376
7.2.3 Terminales multimodo	378
7.2.4 Niveles de interworking	380
7.2.5 Mecanismos de interworking	381
7.3 Interworking entre LTE y otras redes 3GPP.....	388
7.3.1 Marco general	388
7.3.2 Interworking mediante interfaces propias del sistema GPRS	390
7.3.3 Interworking mediante nuevos interfaces entre GPRS y EPC	393
7.3.4 Movilidad entre redes de acceso 3GPP.....	395
7.3.5 Soluciones de interworking para los servicios de voz	401
7.4 Interworking entre LTE y otras redes no 3GPP	404
7.4.1 Marco general	404
7.4.2 Solución genérica para redes Trusted.....	408
7.4.3 Solución genérica para redes Untrusted	415
7.4.4 Movilidad entre sistemas.....	420
7.4.5 Solución de interworking para redes CDMA2000	425
7.5 Referencias Bibliográficas.....	429



I

**ACCESO RADIO MÓVIL:
VISIÓN PERSPECTIVA Y PROSPECTIVA**

1 Acceso radio móvil: visión perspectiva y prospectiva

1.1 Introducción

El sector de las comunicaciones móviles celulares ha mostrado un gran dinamismo en las dos últimas décadas. Los inicios de la década de los 1990 vinieron marcados por el crecimiento exponencial de usuarios de voz al amparo de un entorno cada vez más competitivo y con predominancia de la tecnología GSM como estándar de facto a nivel mundial. Posteriormente, y contrariamente a lo que pronosticaban muchos estudios de mercado, la madurez alcanzada en el servicio de voz no se vio relevada por los servicios de datos en los primeros años de los 2000 de la mano del cambio tecnológico asociado a la implantación del acceso radio WCDMA de UMTS. A nivel global, la principal competencia de UMTS es cdma2000, emanado del 3GPP2, tal y como se ilustra en la Figura 1-1. La clara necesidad de mayores velocidades de transmisión de datos como condición necesaria para el eventual despegue de estos servicios encuentra respuesta en la tecnología HSPA, y equivalentemente EV-DO en el contexto 3GPP2, elementos a la postre facilitadores del crecimiento exponencial del tráfico de datos observado desde 2007, junto con la generalización de las tarifas planas para el acceso a Internet móvil. El camino apuntado por el 3GPP para cubrir las necesidades tecnológicas en el horizonte 2010-2020 tiene a LTE como máximo exponente. La predominancia de LTE supone el fin del camino paralelo del 3GPP2, que abandona el desarrollo de UMB, equivalente a LTE. El contrapunto competitivo para LTE intenta impulsarse desde IEEE con WiMAX 802.16e y posteriormente 802.16m, como solución propiamente IMT-Advanced (sistema 4G), al igual que la propuesta LTE-Advanced por parte del 3GPP.

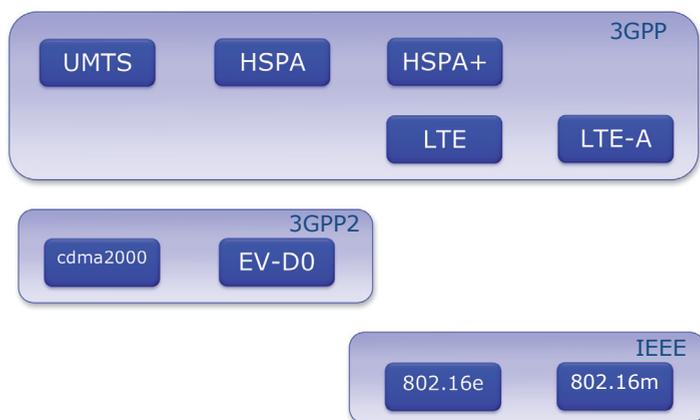


Figura 1-1 Panorámica de sistemas móviles celulares 3G/4G

La evolución tecnológica ilustrada en la Figura 1-1 refleja, sin duda, la existencia de un mercado altamente dinámico, que requiere de una elevada capacidad de adaptación y an-

ticipación para satisfacer las necesidades de los usuarios así como para hacer frente a los competidores, en un entorno marcado también por la evolución de los modelos de negocio, la aparición de nuevos agentes o la modificación del papel jugado por agentes ya presentes. En este sentido, la Figura 1-2 ilustra uno de los principales paradigmas observados en los últimos tiempos. En efecto, mientras que el incremento del tráfico de voz supone un incremento de los ingresos del operador al facturarse por establecimiento de llamada y tiempo de ocupación de los recursos del sistema, en el caso del tráfico de datos se produce un desacople entre ambos, asociado a la adopción de tarifas planas. Por otra parte, tal y como se ilustra en la Figura 1-3, el coste del despliegue y operación de las redes de comunicaciones móviles con la tecnología actual guarda cierta proporcionalidad con el volumen de tráfico a cursar si se pretende ofrecer una QoS adecuada, de manera que para preservar un cierto margen de beneficio es necesario reducir diferentes componentes de la estructura de costes, lo que constituye una de las principales motivaciones para el desarrollo de un sistema como LTE.

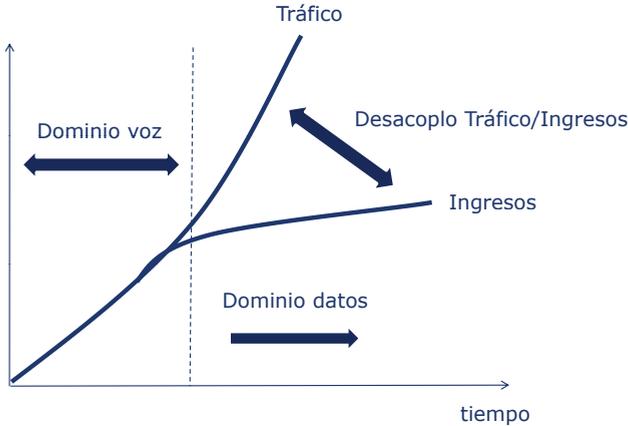


Figura 1-2 Paradigma de la evolución del mercado de comunicaciones móviles

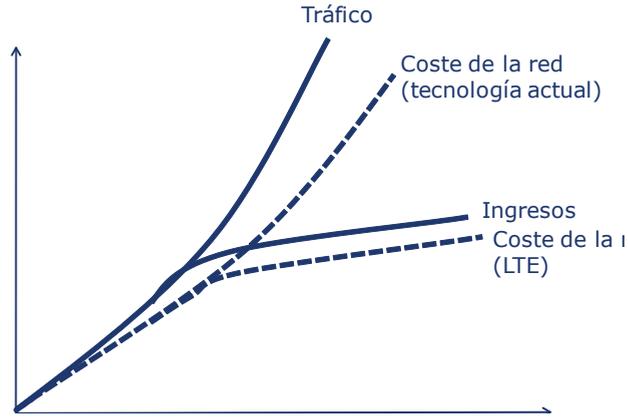


Figura 1-3 Relación del coste de la red con el volumen de tráfico

Con el fin de proporcionar un primer posicionamiento de la tecnología LTE, núcleo de este libro, en un contexto más amplio en el ámbito de los sistemas móviles, este Capítulo pretende describir los principales ingredientes que están involucrados en el proceso de evolución y expansión de las redes de comunicaciones móviles. Como se verá a lo largo de este Capítulo, una visión perspectiva permite ya identificar algunos de estos ingredientes en las tecnologías actualmente en uso, de manera que puede decirse que LTE recoge la evolución de ciertos conceptos bien establecidos y consolidados. Además, la propia evolución tecnológica conlleva la aparición de nuevos fundamentos tecnológicos cuya implantación resulta viable en el horizonte temporal en que se desarrolla LTE. Así pues, puede decirse que LTE como sistema estándar de comunicaciones móviles conjuga evolución y revolución tecnológica, tal y como se describirá con mayor nivel de detalle en los Capítulos posteriores. Por otra parte, y en una visión prospectiva, se identifica también que el posicionamiento final de LTE dependerá también de factores no tecnológicos. Finalmente, este Capítulo describe los aspectos asociados al proceso de estandarización de LTE así como LTE-Advanced.

1.2 Principios de evolución y expansión

Desde la perspectiva de un usuario de un sistema de comunicaciones móviles, éste contrata el acceso a una serie de servicios, con capacidad de movilidad y sobre una cierta extensión geográfica, a los que asocia un cierto nivel de calidad a nivel cualitativo y/o cuantitativo. Desde la perspectiva del operador de la red, éste debe realizar un ejercicio de dimensionado de red con el fin de satisfacer la demanda agregada del conjunto de usuarios del sistema.

A modo ilustrativo, considérese una densidad de usuarios U [usuarios/Km²], cada uno de ellos generando en términos medio un tráfico T [bits/s/usuario]. Para satisfacer esta demanda, considérese que el operador tiene asignado a una estación base dada un ancho de banda B [Hz], que pretende explotar mediante una tecnología caracterizada por una cierta eficiencia espectral E [bits/s/Hz]. Considérese que una estación base tiene capacidad para cubrir una superficie S [Km²]. Así, la capacidad desplegada por el operador es $B \times E / S$ [bits/s/Km²]. El dimensionado adecuado corresponde al caso en que la capacidad desplegada por el operador permite cubrir la demanda de tráfico de los usuarios, esto es, cuando $B \times E / S = U \times T$.

A partir de la igualdad anterior, se desprende que el aumento de la demanda (por la incorporación de un mayor número de usuarios – aumenta U – y/o por el aumento en el uso de las comunicaciones móviles por ejemplo con la incorporación de nuevos servicios –aumenta T –) requiere de una respuesta por parte del operador para satisfacer dicha demanda. El equilibrio podría restablecerse:

- Aumentando B , esto es, accediendo a una mayor cantidad de espectro radioeléctrico.
- Aumentando E , esto es, utilizando una tecnología espectralmente más eficiente.
- Reduciendo S , esto es, desplegando una mayor número de estaciones base para reducir la superficie cubierta por cada una de ellas.

La opción más factible en el corto plazo es reducir S , ya que el aumento de B tiene latencias de orden de magnitud de años debido al impacto regulatorio y el aumento de E tiene en muchos casos impacto a nivel de estandarización, cuyas latencias también son del orden de magnitud de años. No obstante, la movilidad de los usuarios termina por imponer un límite en el tamaño mínimo que puede tener una célula, como se verá con más detalle en el apartado 1.2.1. En esta situación, se diría que la red móvil ha llegado a su límite de capacidad, y los incrementos de demanda deberían satisfacerse a través de cambios regulatorios que permitieran la asignación de mayor cantidad de espectro (aumentar B) y/o el desarrollo y despliegue de nuevas tecnologías que permitieran mejorar la eficiencia espectral ofrecida por las existentes (aumentar E).

Si bien la componente tecnológica, como es el diseño de la interfaz radio en general y de la capa física en particular, suele ser el aspecto más llamativo y al que se le otorga más espectacularidad asociado a la implantación de cada nueva generación de sistemas de comunicaciones móviles, es en realidad la estrategia de despliegue de la red la que proporciona al operador el mayor grado de flexibilidad y la mayor potencialidad en la provisión de aumento de capacidad en la red. En efecto:

- Respecto a la disponibilidad de espectro, partiendo de los 50 MHz iniciales asignados para GSM en la banda de 900 MHz, se añadieron 150 MHz en la banda 1800 MHz y posteriormente 155 MHz en la banda 2100 MHz con la llegada de UMTS. Es decir, en aproximadamente una década se ha conseguido ampliar en un factor aproximadamente 6 la banda disponible. Estableciendo ahora la referencia en los aproximadamente 350 MHz disponibles en 2G/3G, el factor de expansión esperable, puramente en el sentido de aumento de banda disponible para los sistemas de comunicaciones móviles, es limitado (digamos, un factor del orden de 2 en el mejor de los casos). Así, en la componente espectral, las mejoras deben incentivarse a partir de nuevos modelos que conlleven un uso más eficiente del espectro, superando la perspectiva de disponer propiamente de mayor cantidad de espectro.
- Respecto a la evolución de la eficiencia espectral de pico asociada a la tecnología, si tomamos como referencia GSM (9.6 Kbit/s sobre una banda de 200 KHz, esto es, 0.05 bits/s/Hz asociados a un usuario), el paso a UMTS R99 supone un salto hasta 0.4 bits/s/Hz (2 Mbps sobre 5 MHz), mientras que HSDPA proporciona 3 bits/s/Hz (14 Mbit/s en 5 MHz). Esto es, respecto a la tecnología que hoy en día tenemos operativa y que refleja una evolución tecnológica de unos 15 años, se ha introducido un factor de mejora del orden de 60. Estableciendo ahora la referencia en HSDPA, el salto esperable asociado a LTE puede llevar a eficiencias de pico del orden de 5 bits/s/Hz (para el caso 100 Mbit/s sobre 20 MHz) que puede ampliarse a un valor próximo a 16 si se considera MIMO 4x4. Es decir, con respecto a HSDPA, que presenta ya un diseño del enlace radio significativamente sofisticado y eficiente, el factor de mejora esperable con LTE estaría en el rango 2-5 aproximadamente, conseguible mayormente gracias a la introducción de MIMO. Cabe mencionar, no obstante, que la discusión en términos de eficiencia

espectral de pico no debe tomarse de manera dogmática, ya que en la práctica el usuario no experimenta las prestaciones asociadas a la eficiencia espectral de pico sino la eficiencia espectral media.

- Respecto al tamaño de las células, citar que en el Reino Unido (con una superficie del orden de 250,000 Km² y una población de unos 60 millones de habitantes) están desplegadas en 2009, según la aplicación Sitefinder de Ofcom, del orden de 52,000 estaciones base. La capacidad de expansión en este sentido, particularmente considerando el despliegue de femtocélulas, es ciertamente significativa. Por ejemplo, algunas previsiones efectuadas en 2006 [1] pronostican 102 millones de usuarios de femtocélulas sobre 32 millones de puntos de acceso en todo el mundo para 2011.

A partir de estos primeros conceptos básicos, en este Capítulo se elaboran los principios de expansión de la red de acceso radio con el fin de proporcionar una visión panorámica de la evolución de los sistemas de comunicaciones móviles, arrancando en los sistemas 2G y pasando por los sistemas 3G hasta llegar a LTE, núcleo del presente libro, y las puertas de los sistemas 4G.

Según lo comentado anteriormente, el operador de red móvil puede expandir la capacidad de la red a través del despliegue de un mayor número de estaciones base, la disponibilidad de mayor cantidad de espectro o el uso de una tecnología más avanzada. La variable preferida a atacar dependerá tanto de la viabilidad y flexibilidad que ofrezcan cada una de ellas así como de consideraciones económicas al respecto de las mismas. Por ejemplo, si un operador pudiera acceder a mayor cantidad de espectro con un coste de las licencias bajo, ésta podría resultar sin duda la opción más atractiva para el operador. Por el contrario, si resultara que el coste de las licencias para la operación en nuevas bandas frecuenciales fuera muy elevado, podría resultarle más interesante explotar una nueva tecnología sobre las bandas que ya tuviera asignadas, siempre que la regulación lo permitiera.

1.2.1 Despliegue

En términos de despliegue de red, un aumento progresivo de la demanda se cubre con un aumento progresivo de estaciones base. De manera simplificada, podría decirse que, para unos valores de U [usuarios/Km²], T [bits/s/usuario], B [Hz] y E [bits/s/Hz] dados, el área de cobertura de una célula debería ser $S=(B \times E)/(U \times T)$ [Km²]. Antes de pasar a identificar diferentes componentes asociadas al despliegue que se espera jueguen un papel relevante en el contexto de LTE, conviene destacar a continuación el compromiso existente entre cobertura y velocidad de transmisión así como el impacto de la movilidad como elemento limitativo del despliegue.

El compromiso entre cobertura y velocidad de transmisión

La teoría de Shannon [2] establece la capacidad máxima que puede obtenerse en la transmisión desde un elemento transmisor hasta otro receptor a través de un canal de comunicaciones afectado por ruido gaussiano:

$$C = B \times \log_2 \left(1 + \frac{P}{N} \right) = B \times \log_2 \left(1 + \frac{E_b}{N_o} \times \frac{R_b}{B} \right) \geq R_b \quad (1.1)$$

donde B es el ancho de banda y P/N la relación señal a ruido. A su vez, la potencia de señal recibida P se relaciona con la energía del bit E_b a través de la velocidad de transmisión R_b , mientras que la potencia de ruido N se relaciona con la densidad espectral N_o a través del ancho de banda B .

Por otra parte, la potencia recibida P resulta de la potencia empleada en el transmisor P_T una vez ha sufrido las pérdidas de propagación $L(d)$, que dependen de la distancia, d , existente entre transmisor y receptor:

$$E_b = \frac{P_T}{R_b \times L(d)} \quad (1.2)$$

Combinando las expresiones anteriores y expresándolo directamente en escala logarítmica se obtiene:

$$L(d)[dB] \leq P_T[dBm] - N_o[dBm / Hz] - 10 \log \left(\left[2^{R_b/B} - 1 \right] \times B \right) \quad (1.3)$$

Así pues, suponiendo una cierta densidad espectral de potencia de ruido dada N_o y una potencia disponible en el transmisor limitada P_T , las pérdidas máximas de propagación tolerables (y con ello el radio y correspondientemente la superficie cubierta por la célula S [Km²]) se reducen en mayor o menor medida según el ancho de banda disponible B y la velocidad de transmisión R_b de acuerdo con el último término de la expresión anterior.

La movilidad como elemento limitativo en el despliegue

Supóngase un sistema celular regular con células de radio R [m]. A nivel ilustrativo [3], modelemos la distancia que recorre un móvil dentro de una célula como una variable aleatoria con distribución uniforme en el intervalo $[0, 2R]$. El tiempo de permanencia en la célula, τ , si el móvil se desplaza a velocidad uniforme de v [m/s] será también una variable aleatoria. Considerando además que la duración de una conexión, δ , puede modelarse como una variable aleatoria exponencial de media D [s], la probabilidad de que un usuario deba efectuar un handover, P_h , a lo largo de la comunicación vendría dada por [3]:

$$P_h = \int_0^{2R/v} e^{-\tau/D} \frac{1}{2R/v} d\tau = \frac{1}{\alpha} \left[1 - e^{-\alpha} \right] \quad (1.4)$$

donde se ha definido el factor de movilidad, α , como $\alpha = 2 \times R / (v \times D)$. La Figura 1-4 muestra esta función, indicando claramente que:

1. Cuanto más pequeñas sean las células (menor R) más probable será que el usuario deba realizar un *handover*.
2. Cuanto mayor sea la velocidad del móvil (mayor v) más probable será que el usuario deba realizar un *handover*.
3. Cuanto más larga sea la duración media de las comunicaciones (mayor D) más probable será que el usuario deba realizar un *handover*.

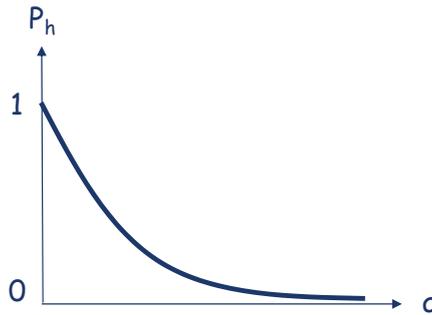


Figura 1-4 Probabilidad de handover versus factor de movilidad

El número medio de *handovers* que efectuará una comunicación, considerando llegadas de Poisson así como el caso ideal en que el sistema proporciona siempre continuidad a la comunicación sería:

$$\bar{N}_h = \sum_{i=1}^{\infty} i \times P_h^{i-1} \times (1 - P_h) = \frac{P_h}{1 - P_h} \quad (1.5)$$

Claramente, la expresión anterior indica que si P_h se aproxima a 1, entonces el número medio de handovers tiende a infinito y el funcionamiento del sistema sería inviable. Esta situación se daría para valores del factor de movilidad tendiendo a 0. Para unas v y D dadas, a medida que se reduce el radio de la célula R aumenta la probabilidad de handover y se reduce el tiempo en que discurre un móvil bajo el área de cobertura de una célula. Puesto que el proceso de señalización asociado al *handover* requiere de un cierto tiempo para su ejecución además de suponer una carga de señalización sobre la red, el aumento del número de procesos de *handover* y la reducción del tiempo entre handovers consecutivos puede hacer inviable la operación práctica de la red.

1.2.1.1 Estructuras multicapa y femtocélulas

En la práctica, la demanda de tráfico no es uniforme ni en la dimensión espacial ni en la temporal. En cuanto a la dimensión temporal, clásicamente se ha utilizado el concepto de "hora cargada", de manera que U y T se toman como los valores esperados de las funciones dependientes del tiempo $U(t)$ y $T(t)$ en el período de tiempo en que la red se ve sometida a la mayor carga (p.e. en horario de oficina en entornos de centro de negocios, en horarios pertinentes en entornos de ocio). En cuanto a la dimensión espacial, la distribución de usuarios presenta fuertes diferencias (p.e. en entornos urbanos frente a entornos rurales), con lo que la visión del despliegue de red conlleva la necesidad de células más pequeñas en los entornos de alto tráfico, ampliándose la superficie cubierta por una célula en entornos rurales.

Así, el despliegue clásico en un entorno urbano hace uso de las denominadas estructuras celulares jerárquicas, en las que se combinan macrocélulas, situadas típicamente en las azoteas de los edificios y proporcionando células de mayor superficie, microcélulas, situadas a pocos metros del suelo con superficies de cobertura más reducidas, y picocélulas, situadas en interiores de edificios y con unos pocos metros de cobertura. De esta manera, las estructuras celulares jerárquicas proporcionan flexibilidad en el despliegue de red y pueden entenderse como inherentemente ligadas a la estrategia de despliegue del operador, con lo que se ha utilizado intensivamente ya desde las primeras fases de GSM. Los trabajos de AT&T Bell Labs constituyen una referencia destacada y pionera en el concepto de microcélulas [4].

Por otra parte, nótese que las estructuras celulares jerárquicas permiten la introducción de mecanismos de gestión de recursos radio que reduzcan el número de handovers que deben realizarse en la red, mediante una apropiada asociación de los usuarios a las distintas capas de células según sea su grado de movilidad. Así por ejemplo, resultaría más adecuado soportar una comunicación desde un coche en movimiento a través de macrocélulas que no a través de microcélulas, ya que en este último caso deberían realizarse *handovers* muy frecuentemente. En definitiva, se trataría de evitar que los usuarios que presenten una v elevada vean reducido su factor de movilidad α mediante la asignación a una célula que presente un mayor R . Una posible manera de estimar la movilidad de los terminales es precisamente a través de la observación de su tasa de *handovers* (esto es, si se observa que una conexión ha requerido varios *handovers* en un breve período de tiempo, puede asumirse un alto grado de movilidad al usuario e intentar traspasar esta comunicación a una macrocélula del entorno).

En el contexto de una red GSM en la que el tráfico es básicamente voz, las microcélulas y picocélulas responden a la necesidad de satisfacer a una fuerte concentración de usuarios, esto es, $S=(B \times E)/(U \times T)$ [Km²] debe reducirse debido a que U es muy elevado. No obstante, con la llegada de HSPA y el uso masivo de acceso a Internet, se incrementa significativamente el volumen de tráfico generado por usuario, con lo que aparece la necesidad de reducir S a causa del incremento de T . La propia evolución de los terminales (por ejemplo, iPhone) y sus funcionalidades conlleva también un aumento de tráfico significativo. Ello deriva en la aparición del concepto de femtocélula, en el que se concibe el despliegue de una estación base de bajo coste por usuario residencial. En este caso, la femtocélula se plantea para proporcionar

un enlace vía radio al usuario desde cualquier ubicación en su entorno doméstico y la propia femtocélula proporciona conectividad a través de una conexión ADSL. Nótese que, al estar la movilidad del usuario confinada en su entorno doméstico la mayor parte del tiempo de utilización del servicio, la problemática de los handovers no aparece como elemento limitativo, aunque S pueda resultar muy reducida. Así, las femtocélulas en el marco de los sistemas 3G avanzados como LTE permitirán que los usuarios puedan gozar de las velocidades de transmisión de pico, sin malgastar capacidad de red para penetrar en interiores, donde se genera gran parte del tráfico a cursar. Otra ventaja que pueden aportar las femtocélulas es que pueden hacer uso de las bandas de frecuencia más altas al tener asociadas coberturas limitadas [5].

1.2.1.2 Redes heterogéneas

El despliegue de LTE comportará la operación de una nueva tecnología, aumentando así el grado de heterogeneidad ya existente hoy en día en los escenarios típicos de comunicaciones móviles. Así, en el caso habitual el despliegue de LTE coexistirá con las redes GERAN y UMTS/HSPA. La explotación de la heterogeneidad de redes de acceso radio requiere lógicamente de la capacidad de inter-operación entre las mismas, tal y como se describirá en el *Capítulo 7: Redes heterogéneas*.

Por otro lado, con el nacimiento de UMTS se despertó un gran interés en el estudio de las ganancias que podrían alcanzarse con la explotación conjunta de los recursos radio correspondiente a redes heterogéneas, a través de los denominados mecanismos de Common Radio Resource Management (CRRM). En este sentido, las estrategias de selección de red y *handover* vertical entre redes GERAN/UMTS permiten conseguir mejoras de capacidad significativas [6][7]. No obstante, cuando las redes heterogéneas incluyen redes de alta velocidad como HSPA y aún en mayor medida LTE, la ganancia de capacidad debida a la explotación conjunta de las redes deja de ser un elemento destacable, ya que la gran diferencia de prestaciones entre las distintas redes hace que la aportación de las redes GERAN/UMTS sea poco significativa. Ello puede entenderse fácilmente a partir del siguiente ejemplo: considérese la función clásica Erlang-B para evaluar la capacidad C_1 soportada por el Sistema #1 que dispone de M_1 canales con una cierta probabilidad de bloqueo, P_B . Considérese el Sistema #2 con M_2 canales y la misma probabilidad de bloqueo, soportando una capacidad C_2 en el mismo escenario. Caso de que este escenario de redes heterogéneas se explotara sin capacidad de interoperación entre las redes, la capacidad total alcanzable sería $C = C_1 + C_2$, con $C_1 = \text{ErlangB}(M_1, P_B)$ y $C_2 = \text{ErlangB}(M_2, P_B)$. Por el contrario, si las redes tuvieran la facilidad de interoperación entre ellas, la capacidad alcanzable sería $C = \text{ErlangB}(M_1 + M_2, P_B)$. La Figura 1-5 muestra la ganancia de capacidad que se consigue con una gestión conjunta de los recursos radio con respecto a una gestión independiente para cada red, esto es, $G = \text{ErlangB}(M_1 + M_2, P_B) / [\text{ErlangB}(M_1, P_B) + \text{ErlangB}(M_2, P_B)]$. Se observa claramente que la ganancia es muy elevada cuando el número de canales es bajo. A medida que aumenta el número de canales (aumenta la capacidad) la ganancia se reduce, limitándose al 10% aproximadamente para los mayores valores mostrados.

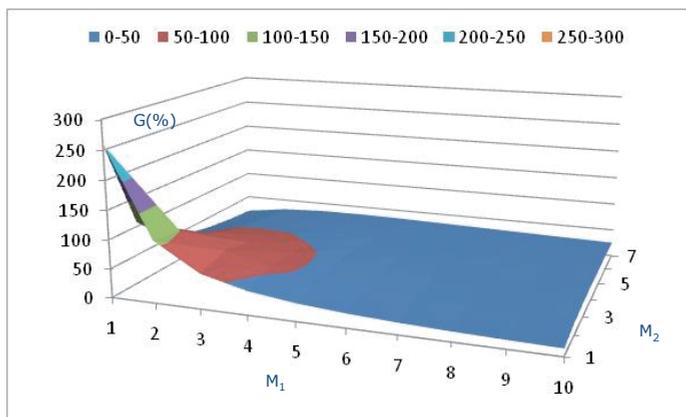


Figura 1-5 Ilustración del denominado “trunking gain”

En cualquier caso, y aunque como se ha ilustrado la ganancia de CRRM en términos de capacidad esperable es discreta, la explotación de la heterogeneidad de las redes sigue siendo un aspecto de gran importancia en el marco del despliegue LTE. En efecto, por una parte las redes *legacy* pueden seguir cursando tráfico asociado a diferentes servicios y así continuar la amortización de las inversiones en estas redes, de manera similar a lo ocurrido en su día con el despliegue de UMTS y la continuidad en el uso de la red GSM para cursar tráfico de voz. Por otro lado, el hecho de que las redes *legacy* ya proporcionen cobertura extensa en el territorio, con disponibilidad ubicua de voz con GSM/UMTS y datos con HSPA, permitirá un despliegue más flexible de los eNodeB de LTE, centrándose al menos inicialmente en la provisión de cobertura en *hot zones*.

1.2.1.3 Mecanismos SON

Con todo lo anterior, en un escenario urbano pueden estar desplegadas un número muy elevado de estaciones base de todo tipo, máxime si se considera un despliegue intensivo de femtocélulas domésticas. En estas condiciones, aparece un nuevo paradigma vinculado a la complejidad de la gestión de una red de acceso radio formada por un número muy elevado de nodos. Así, evoluciona también la concepción de la manera en que una red como LTE deberá operarse y gestionarse, ya que la presión por reducir costes conlleva realizar una explotación de la red mucho más eficiente. En este sentido, los conceptos SON (*Self-Organizing Networks*) se vislumbran como un ingrediente importante, ya que la componente OPEX (*Operational Expenditures*) repercute un peso importante en la estructura de costes.

SON persigue la configuración y optimización automática de la red, existiendo consenso en foros como NGMN de que jugará un papel relevante para hacer el *business case* de LTE más atractivo y, como prueba de ello, SON está recibiendo en los últimos años gran atención. Por ejemplo, en [8] se han formulado los casos de uso en diferentes estadios: planificación (por ejem-

plo, determinar la potencia del eNode B), despliegue (por ejemplo, determinar los parámetros de transporte), optimización (por ejemplo, optimización de parámetros radio bien sea orientada a capacidad, cobertura o prestaciones) y mantenimiento (por ejemplo, actualización de software). El 3GPP también ha reconocido la importancia de SON [9] y el progreso puede verse reflejado por ejemplo en [10], donde se desarrollan casos de uso junto con las funcionalidades requeridas, criterios de evaluación y resultados esperados así como impacto en las especificaciones e interfaces. Ciertamente, el concepto SON es muy ambicioso, con lo que continuará requiriendo esfuerzos de investigación para facilitar su implementación práctica progresiva. Entre los aspectos en los que se debe todavía profundizar pueden citarse (1) resolver adecuadamente el compromiso entre la ganancia en prestaciones obtenibles y la complejidad de implementación en términos de la señalización y medidas necesarias, los requerimientos de computación, etc., (2) asegurar la convergencia de los algoritmos a una solución estable en un tiempo determinado o (3) asegurar la robustez frente a medidas de entrada a los algoritmos corruptas o erróneas.

1.2.1.4 Conceptos SDR en el despliegue de red

Finalmente, en cuanto al despliegue y estructura de la red radio, cabe destacar que la implementación de LTE supondrá la posibilidad de llevar a la práctica algunos de los conceptos que se han ido gestando y madurando en los últimos años en el entorno Software Defined Radio (SDR). Se entiende por SDR un dispositivo radio en el cual algunas de sus funciones de nivel físico (o todas) pueden modificarse por software, lo que proporciona mayor flexibilidad frente al hardware tradicional, permitiendo por ejemplo que se puedan añadir nuevas capacidades o modificar las existentes sin necesidad de reemplazar el hardware. Por otra parte, la enorme evolución en términos de capacidad de computación permite plantear también la separación física de las funcionalidades puramente radio y de procesamiento digital, de manera que la unidad de radio podría estar distribuida en los emplazamientos desplegados sobre todo el territorio mientras que la unidad digital podría ubicarse en un centro de gestión soportando multitud de emplazamientos, unidos por fibra óptica, tal y como ilustra la Figura 1-6. Esta tendencia vendría a tener un cierto paralelismo con la filosofía *cloud computing*, paradigma de los servicios de computación a través de Internet.

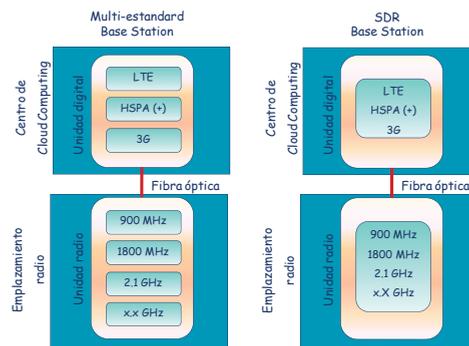


Figura 1-6 Modelo de separación física entre la unidad radio y la unidad digital

Los diferentes aspectos mencionados anteriormente tienen, en definitiva, una misma motivación: conseguir reducir el coste de despliegue y operación de la red para compensar en la medida de lo posible el *gap* cada vez mayor entre el nivel de tráfico que se ofrece a la red, que no cesa de aumentar, y la saturación en el *revenue* que se observa. La Figura 1-7 ilustra a modo de resumen los principales elementos que se espera jueguen un papel destacado en el contexto del despliegue LTE.

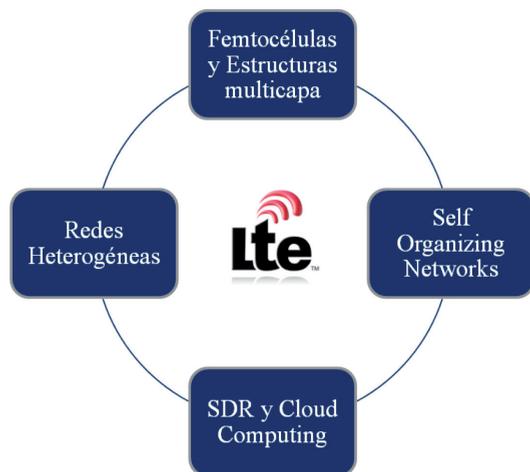


Figura 1-7 Componentes asociadas al despliegue de red en el marco de LTE.

1.2.2 Espectro radioeléctrico

La capacidad de una red puede aumentar de manera directamente proporcional al aumento de ancho de banda que pueda disponerse. Sin embargo, la asignación de espectro resulta en la práctica un proceso altamente complejo, que además lleva asociados unos plazos regulatorios y administrativos que se dilatan a lo largo de varios años, de manera que clásicamente se ha considerado que la variable B del modelo introducido en el apartado 1.1 venía a ser un dato. Así por ejemplo, el espectro asignado inicialmente a GSM correspondía a la banda 890-915 MHz para el enlace de subida y 935-960 MHz para el enlace de bajada.

La asignación de licencias DCS-1800 (1710-1785 MHz y 1805-1880 MHz), en el caso de España realizado en junio de 1998, constituye un claro ejemplo de incremento de capacidad por parte del operador a través del incremento de la banda disponible, B. En este caso, la banda adicional se repartió entre 3 adjudicatarios, al incorporarse Amena (hoy Orange) como operador. Para el caso de explotación DCS-1800 en todos los emplazamientos GSM-900, el tráfico absorbible ($U \times T$) aumenta proporcionalmente al incremento de banda.

En el caso de UMTS, la identificación de bandas de frecuencia se realizó en la WRC'92. La World Radio Conference (WRC) es un evento organizado por la International Telecommunication Union (ITU) -organismo de las Naciones Unidas-, que se suele realizar cada cuatro

años y su objetivo es armonizar la utilización del espectro radioeléctrico a nivel mundial. Las administraciones europeas elaboran sus posiciones técnicas de manera conjunta dentro de la CEPT (European Conference of Postal and Telecommunications Administrations), si bien en última instancia las distintas administraciones europeas negocian en la ITU sobre una base nacional. En particular, se identificaron 155 MHz sobre la banda de 2100 MHz para la componente terrestre de IMT-2000, correspondiendo dos bloques de 60 MHz para el modo FDD y 35 MHz para el modo TDD. Esta componente de espectro IMT-2000 suele denominarse como *core band*. La expectativa de mayor necesidad de espectro para IMT-2000 llevó a la WRC-2000 a identificar las denominadas *new 3G spectrum* en las bandas 2500-2690 MHz.

Por su parte, la WRC'07 identificó espectro adicional para IMT-2000 e IMT-Advanced en las bandas 450-470, 698-806, 2300-2400 y 3400-3600 MHz, aunque la aplicabilidad de las bandas varía a nivel regional y nacional.

No obstante, el aumento continuado del tráfico en las redes móviles, conlleva que deban considerarse todas las componentes posibles para el incremento de la capacidad. En este sentido, en los últimos años se ha planteado la reconsideración de la asignación y uso del espectro radioeléctrico frente a los procedimientos clásicos descritos anteriormente. Las propuestas abarcan múltiples dimensiones, algunas de ellas revolucionarias, que deben ponderarse adecuadamente teniendo en cuenta los aspectos técnicos, regulatorios, impacto en el modelo de negocio y competencia, etc. En las secciones siguientes se presenta una panorámica de estas alternativas.

Una posible clasificación de los modelos de uso del espectro radioeléctrico permite distinguir los siguientes:

1. Modelo de uso exclusivo de derechos. En este caso se realiza la asignación de licencias exclusivas a operadores para ciertas bandas de frecuencias, que pueden estar ligadas o no con el uso de una determinada tecnología y la provisión de determinados servicios.
2. Modelo de uso común del espectro. En este caso se promueve el acceso compartido por parte de una serie de usuarios a una determinada banda del espectro radioeléctrico. La responsabilidad de la gestión de la interferencia en este caso recae directamente entre los propios usuarios. Dentro de este caso puede distinguirse el modelo de uso público (en el que el espectro está abierto a todos los usuarios con los mismos derechos, como sería el caso de la banda ISM – *Industrial, Scientific and Medical*) y el privado (en que es el adjudicatario de una licencia el que permite el acceso al espectro licenciado a un conjunto de usuarios bajo unas condiciones establecidas por el propio adjudicatario).
3. Modelo de uso oportunista del espectro. En este caso se permite que el denominado usuario secundario (que pretende acceder a una banda de espectro licenciada al denominado usuario primario) identifique de manera independiente bandas de espectro que no se estén utilizando en un cierto tiempo y lugar, pudiéndolas utilizar

en ese caso siempre que se asegure que no generará interferencia dañina para el usuario primario. Dentro de este caso puede distinguirse entre el modelo *underlay* (en el que la interferencia al usuario primario se evita transmitiendo una señal con una densidad espectral de potencia que se encuentre por debajo del nivel de ruido, como es el caso de UWB) y el *overlay* (en el que no se da esta coexistencia sino que el usuario secundario debe detectar y aprovechar los huecos espectrales libres de usuario primario).

La tendencia es pues, de manera clara, hacia un uso más flexible y dinámico del espectro radioeléctrico.

1.2.2.1 Refarming

Las bandas de 900 y 1800 MHz quedaron restringidas regulatoriamente a través de la Directiva 87/372/ECC [11] conocida como "Directiva GSM" a servicios de telefonía móvil GSM. Posteriormente, el sistema UMTS se despliega en la banda de 2100 MHz, siguiendo el mismo modelo de uso exclusivo de derechos (mediante otorgación de licencias). Dentro del propio modelo de uso exclusivo de derechos, *spectrum refarming* se refiere a la posibilidad de explotar la tecnología más conveniente en la banda más conveniente dentro de las asignadas en licencia a un operador, lo que supone un primer principio de flexibilidad espectral. La CEPT define *refarming* como "una combinación de medidas administrativas, financieras y técnicas, presentes y futuras, dentro de los límites de regulación de frecuencias, con el objeto de hacer una banda de frecuencia específica disponible a otro tipo de uso o tecnología. Estas medidas pueden ser a corto, medio o largo plazo". Así, el *refarming* requiere de una modificación regulatoria en este caso, para permitir el uso de las bandas de 900 MHz y 1800 MHz bajo los principios de neutralidad tecnológica y de servicios, los cuales son objetivos de la política común europea. A finales de 2008 la Comisión Europea presentó una propuesta para la revisión de la "Directiva GSM" en esta dirección. La propuesta fue aprobada por el Parlamento en mayo de 2009, y publicada en Diario Oficial de la Unión Europea el 20 de octubre de 2009 [12] para su aplicación a partir de mayo de 2010. En cualquier caso, el debate del *refarming* de GSM se está ya llevando a cabo en los diversos países de la Unión Europea, dadas las fuertes implicaciones asociadas en términos de competencia y desarrollo del sector.

1.2.2.2 Canalización flexible

La canalización de un sistema de comunicaciones móviles es uno de los parámetros fundamentales que lo caracterizan y su elección está condicionada por múltiples factores. La canalización está directamente relacionada con la velocidad de transmisión alcanzable, de manera que para alcanzar mayores velocidades de transmisión se requiere aumentar la canalización. Por otro lado, y a partir de la teoría de muestreo de Nyquist, la canalización tiene una incidencia directa en la complejidad de implementación del sistema (por ejemplo, el

convertor A/D) y en la complejidad computacional asociada al tratamiento de señal. Así pues, al concebirse una nueva generación de comunicaciones móviles pensando en un determinado horizonte temporal para su implementación práctica, se establecen sus capacidades de transmisión y se selecciona una cierta canalización de acuerdo con el estado del arte esperable en los ámbitos de hardware y procesado en la época de su implementación.

En GSM la canalización se estableció en 200 KHz, valor asumible con la tecnología disponible a principios de la década de 1990. En UMTS, y aprovechando los avances de toda una década, se escogieron 5 MHz.

En el caso de LTE, para alcanzar velocidades de transmisión del orden de 100 Mbits/s en el *downlink* se considera una canalización de 20 MHz. Para LTE-Advanced, en el que se pretende llegar a 1 Gbit/s, la canalización se aumenta hasta 100 MHz.

Tal y como se ha comentado anteriormente, se observa una notable disparidad en las bandas de frecuencias y las posibilidades espectrales sobre las que desplegar LTE, de manera que sería deseable que LTE ofreciera una elevada flexibilidad en este sentido. Por ejemplo, en el caso de considerar *refarming* sobre la banda de 2100 MHz, un operador determinado puede disponer de bloques de 5 MHz de ancho de banda inicialmente asociados a portadoras UMTS, con lo que sería interesante que LTE tuviera la capacidad de operar también con esta canalización. Obviamente, las prestaciones alcanzables con LTE operando sobre un ancho de banda de 5 MHz quedarían aproximadamente escaladas en un factor 4, correspondiente a la reducción de canalización con respecto a los 20MHz. La canalización flexible es, por tanto, un concepto de gran interés cuando se plantea la migración gradual hacia LTE de bandas espectrales actualmente ocupadas por 2G/3G.

En particular, las canalizaciones posibles para LTE son 1.4 MHz, 3 MHz, 5 MHz, 10 MHz, 15 MHz y 20 MHz. Citar también que, bajo el concepto de flexibilidad espectral, se incluye también la posibilidad de operación tanto sobre bandas pareadas como no pareadas, de manera que LTE permite también su explotación tanto en FDD como TDD, incrementando así la flexibilidad en su despliegue.

1.2.2.3 Dividendo digital

La Comisión Europea ha defendido que Europa debía perseguir la meta común de maximizar el potencial social y económico del "dividendo digital" derivado de las bandas de espectro potencialmente libres tras el cambio a difusión digital [13]. El dividendo digital se deriva de la capacidad de transmitir hasta 8 canales de TV digital en el espectro que anteriormente ocupaba un único canal de TV analógica (la ganancia puede incluso ser mayor si se utilizan estándares más avanzados como DVB-T2 para la infraestructura y MPEG-4 para la compresión).

Todos los Estados miembros deben completar el abandono de la televisión analógica para 2012 a más tardar, aunque algunos países como Alemania, Finlandia o Suecia ya lo completaron antes de fin de 2009. La Comisión Europea también ha enfatizado la necesidad y posibles

aproximaciones para conseguir una apropiada coordinación a nivel de la Unión Europea [14]. La necesidad de coordinación también ha sido reconocida en diversas ocasiones por el Radio Spectrum Policy Group (RSPG), que aconseja a la Comisión Europea en el desarrollo de la política de espectro radioeléctrico. Para comprender el posible impacto social y económico del dividendo digital y así poder derivar la estrategia más apropiada desde la perspectiva global de la Unión Europea, la Comisión llevó a cabo un amplio estudio [15], cuyos resultados se presentaron públicamente el 9 de septiembre de 2009 en Bruselas. Según el mismo, si antes de 2015 se consigue una coordinación europea adecuada del espectro del dividendo digital en su conjunto, su impacto económico potencial se incrementaría entre 20.000 y 50.000 millones de euros adicionales a lo largo de 15 años, dependiendo principalmente del nivel real de la futura demanda de servicios tales como la radiodifusión terrestre avanzada y la banda ancha inalámbrica.

Por otra parte, el 10 de julio de 2009, la Comisión Europea lanzó una consulta pública sobre el dividendo digital, con el objetivo de recibir comentarios y puntos de vista de los diferentes *stakeholders* [16]. La consulta incluyó un punto sobre la conveniencia de poner disponible la banda de 800 MHz para redes de comunicaciones electrónicas de baja/media potencia bajo condiciones técnicas armonizadas y siguiendo los principios de neutralidad tecnológica y de servicio. De dicha consulta se deriva un alto nivel de consenso en cuanto a establecer un modelo para abrir de manera coherente la subbanda de 790-862 MHz a los servicios de comunicaciones electrónicas adoptando unas condiciones técnicas de uso armonizadas.

Sobre las bases anteriores, la Comisión Europea adoptó el 28 de octubre de 2009 una Recomendación [17] y una Comunicación [18] al Parlamento y al Consejo con una serie de propuestas relativas a un enfoque común con respecto al dividendo digital en Europa, de manera que pueda avanzarse de inmediato en lo que se refiere a los retos más urgentes. Este movimiento responde también al hecho de que varios Estados miembros están ya, o lo estarán en breve, en situación de decidir la apertura de la subbanda de 790-862 MHz para servicios de comunicaciones electrónicas. Dado que los Estados miembros se ven sometidos a una presión creciente para actuar, si no se cuenta urgentemente con un modelo comunitario de enfoque coordinado se corre el riesgo de fragmentación del mercado interior y de pérdida de las economías de escala que podrían obtenerse. Literalmente, la Recomendación se centra en dos aspectos:

1. Que los Estados miembros adopten todas las medidas necesarias para garantizar que todos los servicios de radiodifusión televisiva terrestre utilicen la tecnología de transmisión digital y abandonen la tecnología de transmisión analógica en su territorio antes del 1 de enero de 2012.
2. Que los Estados miembros apoyen la labor reguladora tendente a establecer en la Comunidad unas condiciones de uso armonizadas de la subbanda de 790-862 MHz para servicios de comunicaciones electrónicas distintos y adicionales a los servicios de radiodifusión y se abstengan de adoptar cualquier medida que pueda obstaculizar o impedir el despliegue de dichos servicios de comunicaciones en la mencionada subbanda.

En este contexto, a nivel europeo destaca el papel activo del Reino Unido, a través del regulador Ofcom. Ofcom ya lanzó a finales de 2005 el proyecto "Digital Dividend Review (DDR)" para examinar las distintas opciones asociadas al dividendo digital y posibles nuevos usos del espectro disponible [19]. En todo caso, la explotación del dividendo digital en Europa se materializará con un retraso significativo respecto a Estados Unidos, donde la transición a la TV digital ya está completada y la mayor parte del espectro liberado ya ha sido subastado, asignándose principalmente a los operadores móviles para el despliegue de LTE en la banda de 700 MHz.

1.2.2.4 Cognitive radio

El crecimiento explosivo del uso de bandas no licenciadas (por ejemplo WiFi en ISM) y los avances tecnológicos que ello ha motivado, llevaron a la FCC a considerar la apertura de nuevas bandas bajo esta modalidad [20]. Al mismo tiempo, medidas de ocupación de espectro mostraron que las bandas licenciadas como las de TV están poco utilizadas [21]. *Cognitive radio* aparece como una solución al bajo uso del espectro, ya que permite un uso del mismo de manera flexible, eficiente y fiable a partir de la capacidad de adaptar en tiempo real las características de operación a las condiciones del entorno. Los dispositivos *cognitive radio* podrían explotar de manera inteligente amplias bandas de espectro inutilizadas sin interferir a dispositivos que pudieran tener licencia de operación en estas u otras bandas frecuenciales [22].

El concepto genérico *cognitive radio* inició una primera vía de materialización a partir de la "TV band Notice of Proposed Rule Making" de la FCC en mayo de 2004 [23], que propone permitir la operación de dispositivos no licenciados en las bandas de TV siempre y cuando se garantice que no se producirá interferencia dañina a los servicios licenciados (por ejemplo, un receptor de TV). Motivado por la NPRM, en noviembre de 2004 se formó el grupo de trabajo IEEE 802.22 con el objetivo de definir una interfaz radio basada en *cognitive radio* para la operación no licenciada en bandas de TV.

Por otra parte, la consulta pública de la Comisión Europea sobre el dividendo digital [16] incluyó también un punto referente al uso de equipos *cognitive radio* sobre la base de unos requisitos técnicos comunes en Europa y la potencial consideración de un marco regulatorio coordinado para el tratamiento de los *usage rights* en los *white spaces*. De nuevo, en Europa destaca el papel activo de Ofcom. A finales de 2007 [19], Ofcom aboga por permitir el uso de dispositivos *cognitive radio*, siempre y cuando quede garantizado que no introducen interferencia dañina a los usuarios licenciados. Continuando con su trabajo de anticipación en este ámbito, Ofcom dio un paso más al proponer una serie de parámetros técnicos para la correcta operación de usuarios *cognitive* no licenciados [24], que finalmente ha derivado en [25]. Como conclusiones destacadas puede citarse que, si bien la operación podría basarse exclusivamente en *sensing*, la aportación de geo-localización se considera en general muy importante e incluso esencial, sobre todo al considerar las dificultades de implementación de los mecanismos de *sensing*. En todo caso, el documento reconoce la necesidad de más estudios al respecto.

Un escenario que podría vislumbrarse como plausible es la explotación de los principios de *cognitive radio* en femtocélulas. La relevancia de las femtocélulas ya se ha descrito en el apartado 1.2.1.1. En una visión más avanzada, una femtocélula podría operar no sólo en las bandas licenciadas al operador sino también en otras siguiendo los principios de *cognitive radio*, ya que la operación en interiores facilitaría el aislamiento radioeléctrico necesario.

A nivel español, la "Consulta pública sobre el uso de la banda de frecuencias de 2.500 a 2.690 MHz y sobre posibles nuevas modalidades de explotación de las bandas de frecuencias de 900 MHz, 1.800 MHz y 3,5 GHz" impulsada por la Secretaría de Estado de Telecomunicaciones y para la Sociedad de la Información llevada a cabo en verano de 2008 es un indicio de la importancia que están finalmente tomando estos aspectos, si bien en dicha consulta no se incluyó el debate del dividendo digital. En este caso, los resultados de la consulta reforzaron la conveniencia de los principios de neutralidad tecnológica y el *refarming*. En cuanto a la banda de 2.6 GHz, en general las respuestas recibidas no consideraban urgente la licitación de la misma. En este último aspecto, resulta razonable adoptar una posición prudente, ya que puede resultar complicado efectuar una estimación del valor económico de dicha banda si previamente no está clarificado el escenario resultante del *refarming* y el dividendo digital.

Los diferentes aspectos mencionados anteriormente tienen, en definitiva, una misma motivación: conseguir un uso más eficiente y flexible del espectro radioeléctrico. La Figura 1-8 ilustra a modo de resumen los principales elementos que se espera jueguen un papel destacado en el contexto del despliegue LTE.

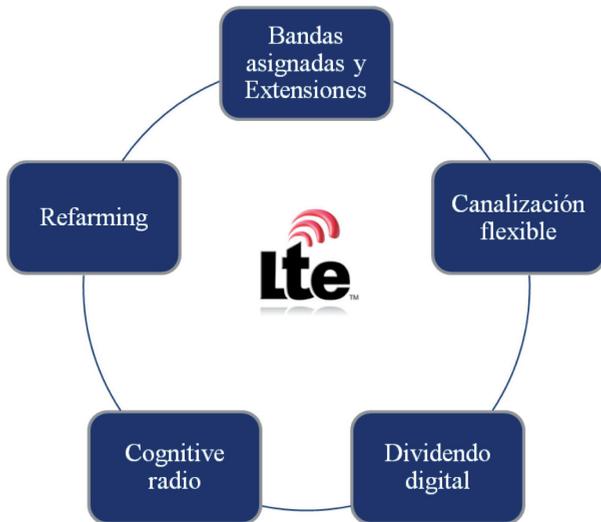


Figura 1-8 Componentes asociadas al espectro radioeléctrico en el marco de LTE

1.2.3 Tecnología y eficiencia espectral

Aunque el concepto de eficiencia espectral E [bits/s/Hz] resulta a nivel teórico un indicador claro y evidente de las prestaciones que puede ofrecer un sistema de comunicaciones móviles y, con ello, puede constituirse como base de comparación para la bondad de diferentes soluciones, en la práctica resulta mucho más complicado, ya que depende de múltiples elementos y no sólo del diseño de la interfaz radio (técnica de acceso múltiple, técnicas de ingeniería radio, etc.) sino también de su explotación práctica en un entorno con múltiples estaciones base.

A principios de los años 90, el encendido debate de si la 3G debía basarse en un acceso TDMA o en un acceso CDMA se planteó en muchas ocasiones en términos comparativos de eficiencia espectral. Célebres trabajos de los fundadores de Qualcomm como [26] pronosticaban que un acceso CDMA ofrecía una capacidad muy superior a cualquier otro esquema de acceso propuesto, con lo que el trabajo concluía que CDMA debía ser la elección lógica para todas las aplicaciones de telefonía celular. No obstante, llevados a la realidad práctica la ganancia no es ni mucho menos tan significativa, y estudios posteriores más detallados concluían que CDMA y TDMA ofrecían eficiencias espectrales similares [27]. Cabe señalar que la sensibilidad de una técnica de acceso u otra a una misma imperfección en la implementación de un determinado subsistema radio o a un determinado efecto del canal móvil puede ser muy diferente. Así por ejemplo, considerar un control de potencia ideal en un análisis comparativo tendería a sobrevalorar el comportamiento de un acceso CDMA, ya que éste es un aspecto crítico en el funcionamiento del mismo. En este sentido, y desde el punto de vista de la elección de la técnica de acceso radio para la 4G, debe decirse que la atención no se ha centrado tanto en la mejora que puede ofrecer un acceso OFDMA per se frente a otras posibles alternativas.

Sea como fuere, puede afirmarse que, conceptualmente, el principal objetivo de la evolución tecnológica en el ámbito de las comunicaciones móviles es el de desarrollar interfaces radio que mejoren en la mayor medida posible la eficiencia espectral. En este sentido, la teoría de Shannon marca una cota del máximo rendimiento que puede obtenerse en un enlace radio entre un transmisor y un receptor y, por lo tanto, permite valorar el grado de bondad de un diseño determinado.

En este apartado se pretende proporcionar una visión panorámica, aunque no completamente exhaustiva, de distintos aspectos que se han ido introduciendo históricamente en el ámbito de las comunicaciones móviles con el fin de ir mejorando su rendimiento espectral. Esta visión panorámica permitirá concluir que algunos de los ingredientes básicos para conseguir un diseño LTE ya muy próximo a la cota de Shannon, resultan de la evolución de conceptos ya presentes en los sistemas 2G (GSM/GPRS) y/o 3G (UMTS/HSPA). Junto a estos principios ya establecidos en el pasado, LTE se fundamenta también sobre nuevos conceptos facilitados por el propio progreso tecnológico.

1.2.3.1 Transmisión y acceso múltiple multiportadora

El diseño de la interfaz radio es, obviamente, uno de los aspectos más relevantes y definitorios de un sistema de comunicaciones móviles. En el caso de GSM, se diseñó una interfaz radio basado en TDMA ajustado a las características del servicio de voz y del codificador estandarizado (RPE-LTP a 13 Kbits/s). En el caso de UMTS, y sobre la base de una potencial mayor eficiencia así como mayor flexibilidad para poder acomodar servicios multimedia, la solución adoptada fue WCDMA. En el caso de LTE (y posteriormente LTE-A), el hecho de considerarse anchos de banda de hasta 20 MHz (y hasta 100 MHz para LTE-A) requirió de nuevo considerar los principios de transmisión y acceso múltiple que pudieran resultar más apropiados.

La transmisión de una señal sobre 20 MHz de ancho de banda estará afectada por un comportamiento del canal móvil que introducirá una fuerte distorsión por interferencia intersimbólica derivada de la propagación multicamino [3]. En estas condiciones, la consideración de la técnica de transmisión OFDM (Orthogonal Frequency Division Multiplexing), consistente en multiplexar un conjunto de símbolos sobre un conjunto de subportadoras ortogonales entre sí, permite eliminar los efectos de la interferencia intersimbólica gracias a la introducción del denominado prefijo cíclico, tal y como se explicará en detalle en el *Capítulo 4: Tecnologías de nivel físico*. El uso de OFDM permite evitar la transmisión de una única señal de banda ancha mediante la transmisión de un conjunto de señales de banda estrecha ortogonales entre sí, resultando por tanto más robusta frente a la propagación multicamino.

Sobre la base de una transmisión OFDM, el empleo de la técnica de acceso múltiple OFDMA proporciona un alto grado de flexibilidad al permitir la asignación dinámica de las diferentes subportadoras a los diferentes usuarios, tal y como se verá en detalle tanto en el *Capítulo 4: Tecnologías de nivel físico* como en el *Capítulo 6: Gestión de recursos radio y del espectro radioeléctrico*. También se describirá en detalle la técnica SC-FDMA, empleada en el enlace ascendente para mitigar los inconvenientes que presenta OFDMA en cuanto a que la potencia instantánea transmitida puede ser significativamente superior a la potencia media, lo que dificulta la realización del amplificador de potencia a incorporar en el terminal móvil.

Finalmente, cabe mencionar que la elección del esquema de transmisión y acceso múltiple viene también condicionada por la complejidad y coste de implementación (incluyendo aquí los aspectos de propiedad intelectual) así como por las propias dinámicas de la industria de fabricación que impulsa de manera natural a la renovación tecnológica. En este caso, el diseño del acceso radio LTE facilita su implementación en el dominio digital a través de la realización de los procesos FFT/IFFT de manera eficiente.

1.2.3.2 Modulación y codificación adaptativas

La estructura de ráfaga GSM está compuesta por un total de 148 bits, que finalmente se transmiten sobre el canal radio con una modulación binaria GMSK. Parte de los bits se dedi-

can a procedimientos de nivel físico como la sincronización, ecualización, etc., otra parte son propiamente datos de usuario y otra parte se dedica a incorporar redundancia mediante un proceso de codificación de canal para permitir la corrección de errores en recepción. En el caso de GSM, se define un único esquema de codificación de canal para la transmisión de la información de voz, y se aplica dicho esquema a todas y cada una de las ráfagas transmitidas por/para todos y cada uno de los usuarios del sistema.

Mientras que en muchas situaciones prácticas la inclusión de los bits de redundancia resulta altamente beneficiosa para mejorar la fiabilidad del enlace, también es cierto que en otras ocasiones (por ejemplo un usuario que se encuentre muy próximo a la estación base y con visión directa a la misma) esta redundancia no sería necesaria. En estos últimos casos, si pudieran transmitirse bits de información de usuario en lugar de bits de redundancia se conseguiría mejorar la eficiencia espectral del sistema. Esta mejora ya se incorpora en GPRS, que introduce 4 esquemas de codificación de canal, con el fin de poder aprovechar las condiciones favorables. El concepto anterior puede también extenderse a la modulación del canal, de manera que EGPRS permite la transmisión bien en GMSK (1 bit por símbolo de canal) bien en 8-PSK (3 bits por símbolo de canal) [28]. De acuerdo con esto, en EGPRS se definen hasta 9 esquemas de modulación y codificación (MCS). Así por ejemplo, MCS-1 utiliza GMSK y codificación de tasa 0.53, MCS-3 utiliza tasa de código 0.8 y MCS-9 utiliza 8-PSK y tasa de código 1.0 (es decir, sin redundancia).

La Figura 1-9 muestra el *throughput* alcanzable para los distintos MCS en función de la relación señal a interferencia (C/I). De su observación se desprende que si el sistema tuviera la capacidad de adaptar el MCS utilizado para la transmisión en una determinada conexión a las condiciones dinámicas que fuera observando dicha conexión (variaciones de la C/I), se conseguiría mejorar el rendimiento del sistema. Esta adaptación permitiría obtener unas prestaciones que vendrían a ser la envolvente de las prestaciones obtenibles con los diferentes MCS.

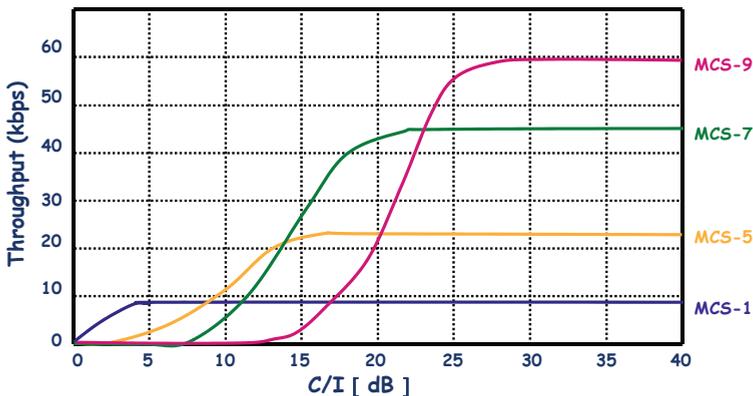


Figura 1-9 *Throughput* obtenido para diferentes MCS en distintas condiciones de C/I

En el caso de HSDPA, se recogen y amplían los conceptos anteriores. Por ejemplo, HSDPA permite el uso de modulación 16-QAM [29]. Por otra parte, en UMTS se recogen también los avances en el campo de la codificación de canal y se estandariza el uso de turbo-códigos, propuestos por primera vez en 1993 [30].

En el caso de LTE, se recogen los conceptos anteriores y se mejoran sus capacidades. Así por ejemplo, el proceso de turbo-codificación en LTE introduce algunas modificaciones con respecto a HSPA, como es la incorporación de un entrelazado QPP (Quadrature Permutation Polynomial), que permite paralelizar la decodificación sin riesgo de contención cuando los distintos procesos paralelos acceden a la memoria del entrelazado [31]. En cuanto a la modulación, LTE considera QPSK, 16-QAM y se amplía también a la posibilidad de utilizar 64-QAM. Más información relativa a estos aspectos puede encontrarse en el *Capítulo 5: Interfaz radio del sistema LTE*.

1.2.3.3 Transmisión por paquetes y packet scheduling

En GSM la transmisión de voz sobre la interfaz radio se lleva a cabo en modo circuito, ya que a un usuario determinado se le asignan unos recursos fijos (en este caso la transmisión de un total de 24 ráfagas en un período de 120 ms) desde el inicio hasta el fin de la comunicación [32]. La transmisión en modo circuito resulta apropiada en el caso en que las fuentes de tráfico generen información de manera continuada, ya que así se aprovechan los recursos asignados. En el caso de que las fuentes de tráfico sean discontinuas la operación en modo circuito resulta en un desaprovechamiento de recursos, ya que un usuario determinado puede desperdiciar la oportunidad de transmisión debido a que no tiene información que transmitir a la vez que otro usuario puede tener tráfico para inyectar al canal pero no dispone de recursos para la transmisión.

El acceso a Internet constituye un claro ejemplo de tráfico a ráfagas, con lo que se hace necesario incorporar otra estrategia para sustentar la comunicación, en este caso la transmisión en modo paquete. GPRS supone el primer exponente en la transmisión en modo paquete y su implantación a finales de los años 90 viene motivada por las expectativas en aquel momento de la explosión del acceso a Internet a través del móvil. En una transmisión por paquetes la asignación de recursos a usuarios es flexible y, por tanto, tiene la capacidad de adaptarse a las necesidades de cada momento. Ello es posible gracias a la sofisticación de los mecanismos de gestión de recursos radio, que incorporan una nueva funcionalidad denominada *packet scheduling* que se encarga de decidir la asignación dinámica a usuarios. En el caso de GPRS, esta funcionalidad se implementa en el BSC. Claramente, la transmisión por paquetes con un algoritmo de *packet scheduling* adecuado consigue mejorar la eficiencia espectral.

Así pues, la transmisión por paquetes en la interfaz radio y el mecanismo de *packet scheduling* están también presentes en UMTS desde su primera versión (Release 99) con la estandarización del canal DSCH (*Downlink Shared Channel*), que opera sobre la base de una trama UMTS de 10 ms y con un *packet scheduling* desde el RNC. Cabe decir que el canal DSCH no

llegó a explotarse en la práctica, ya que ante las expectativas de implantación de HSDPA los operadores prefirieron esperar a HSDPA para introducir estos conceptos sobre las redes 3G.

En el contexto 3G, HSDPA supone la incorporación de la transmisión en modo paquete sobre la interfaz radio, permitiendo en su primera implantación una velocidad teórica máxima de 14.4 Mbits/s (conseguible con 15 códigos OVSF en paralelo, modulación 16-QAM y sin codificación de canal). Nótese que el aumento de la velocidad de transmisión sobre la interfaz radio conlleva la conveniencia de reducir la referencias temporales del sistema (por ejemplo, la periodicidad con la que el algoritmo de *packet scheduling* toma decisiones de asignación de recursos a los distintos usuarios) para evitar una pérdida de eficiencia. Así por ejemplo, en HSDPA se introduce una duración de la trama de 2 ms frente a los 10 ms de UMTS Release 99.

Cabe destacar que el algoritmo de *packet scheduling* en sí mismo influye en la eficiencia espectral obtenida, ya que según lo acertado de sus decisiones se logrará un mayor o menor aprovechamiento de los recursos radio. En este sentido, la mejor granularidad temporal del sistema permite que el algoritmo de *packet scheduling* pueda asignar los recursos no sólo a los usuarios que tienen información a transmitir sino que, entre estos, a aquellos que tengan condiciones más favorables en su canal radio particular en un momento dado. En particular, típicamente el algoritmo de *packet scheduling* en HSDPA toma sus decisiones considerando el llamado CQI (*Channel Quality Indicator*), reportado por los móviles.

La relevancia de la transmisión en modo paquete queda evidenciada en LTE, ya que LTE no proporciona canales dedicados sino únicamente un canal compartido en *downlink* y otro en *uplink*. Ello confiere mayor importancia al mecanismo de *packet scheduling*, sobre el que recae la responsabilidad de asignar los recursos radio de manera que pueda satisfacerse la QoS establecida para cada conexión. Los aspectos de la configuración física de los canales de transmisión por paquetes en LTE se elaboran en el *Capítulo 5 – Interfaz radio del sistema LTE*, mientras que los aspectos de gestión se tratan en el *Capítulo 6: Gestión de recursos radio y del espectro radioeléctrico*.

1.2.3.4 Arquitectura de la red

La arquitectura de la red de acceso radio (RAN – Radio Access Network) ha sido tradicionalmente jerárquica, formada por un conjunto de estaciones base (BTS en GSM, NodeB en UMTS) conectadas a un nodo controlador (BSC en GSM, RNC en UMTS). Tanto en GSM como en UMTS *Release 99*, las estaciones base realizan funciones de transmisión y recepción de las señales, pero no se le atribuyen mecanismos de gestión, decisión y control.

La evolución de las funcionalidades y capacidades de la interfaz radio tiene implicaciones sobre la arquitectura de la red, requiriendo también una evolución en la misma. Por ejemplo, en el caso de HSDPA, el algoritmo de *packet scheduling* pasa a implementarse en el NodeB en lugar de hacerlo en la RNC con el fin de reducir el *round trip delay* (se evita el retardo asociado a la interfaz lub entre NodeB y RNC) y así permitir una mayor efectividad del algoritmo

de *packet scheduling* dependiente de las condiciones radio. La implementación del *packet scheduling* en el NodeB en HSDPA marca un primer hito en la descentralización de funciones en la RAN, tal y como muestra la Figura 1-10. En el caso (a) se muestra R99 y en (b) HSDPA.

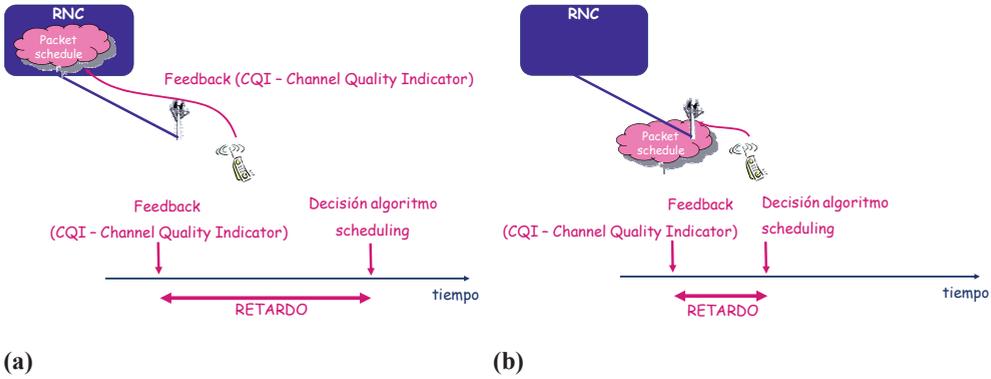


Figura 1-10 Reducción de retardo asociada al *packet scheduler* en HSDPA

Extendiendo el camino iniciado por HSDPA, la estación base de E-UTRAN, denominada eNodeB, ejecuta todos los protocolos de la interfaz radio, tanto del plano de usuario como del plano de control, y lleva a cabo toda la gestión de la interfaz radio. Además del mencionado *packet scheduling*, la interfaz radio LTE incorpora también mecanismos H-ARQ (véase *Capítulo 5 – Interfaz radio del sistema LTE*), que combinan la corrección de errores en recepción a través de la incorporación de redundancia (FEC) con la retransmisión de paquetes (ARQ) a partir de códigos detectores de errores. Los mecanismos de *soft-combining* almacenan los paquetes erróneos (que aunque contengan errores también contienen información útil) en el receptor para combinarlos con los paquetes retransmitidos para dar lugar a un único paquete resultante con una fiabilidad mejor que cada uno de los paquetes constituyentes. A su vez, estos mecanismos pueden aplicar *incremental redundancy* (IR) de manera que la tasa de codificación (número de bits de información / número total de bits transmitidos) se puede ir variando para obtener un mayor grado de protección de un paquete a medida que se efectúan retransmisiones. De nuevo, estos principios ya se explotaban en HSDPA, con lo que LTE básicamente recoge estas capacidades, que se ven mejoradas por el hecho de que toda la gestión se pasa a llevar a cabo desde el eNodeB.

Por otra parte, cabe mencionar que la distribución de funciones a los nodos extremos de la red se ha visto facilitada por el incremento significativo de capacidad de computación de los equipos a costes reducidos, de manera que el cambio esencial es que E-UTRAN sigue una arquitectura “plana”, sin ningún nivel de jerarquización, tal y como se describe en detalle en el *Capítulo 2: Arquitectura del sistema LTE*. Nótese también que la gestión eficiente y robusta del enlace radio LTE, al igual que ya ocurre con HSPA, hace que no sean necesarios mecanismos de diversidad adicionales como el *soft-handover* de WCDMA (que requiere de

funcionalidades de control y enrutado del tráfico asociadas a la RNC), lo que facilita la adopción de una arquitectura plana.

1.2.3.5 Técnicas multi-antena

Las técnicas multi-antena (MIMO) se refieren en general a estructuras de transmisión y/o recepción que utilizan diversas antenas acompañadas de técnicas de procesamiento de señal con distintos niveles de complejidad. Las técnicas multi-antena permiten mejorar las prestaciones del sistema, bien sea aumentando la capacidad (más usuarios por célula), mejorando la cobertura o mejorando la velocidad de transmisión por usuario. El *Capítulo 4: Tecnologías de nivel físico* proporciona una visión general de los principios teóricos sobre los que se sustentan estas técnicas, desarrolladas intensamente desde la segunda mitad de los años 90, en que se empiezan a formular las primeras ideas al respecto [33]. Mientras que este tipo de estructuras multi-antena no resultaban todavía viables para la implementación práctica en el momento en que se estandarizó UMTS, sí resultan posibles para su incorporación tanto en HSPA+ como en LTE, suponiendo un salto cuantitativo importante en las cotas de eficiencia espectral alcanzables.

Los diferentes aspectos mencionados anteriormente tienen, en definitiva, una misma motivación: conseguir el diseño y explotación de una interfaz radio lo más eficiente posible, soportado por una red de acceso que presente también la arquitectura más apropiada. La Figura 1-11 ilustra a modo de resumen los principales elementos que juegan un papel destacado en el contexto de la interfaz radio LTE. Nótese que, a diferencia de lo descrito en el apartado 1.2.1 y el apartado 1.2.2 en los que existe cierto nivel de indeterminación del papel a jugar por los diferentes ingredientes (ya que dependen de aspectos como la regulación, las estrategias del operador y el mercado mismo), las componentes tecnológicas de LTE están ya establecidas a partir de la definición del propio estándar.



Figura 1-11 Componentes destacados de la interfaz radio LTE

1.3 Estandarización

Las tecnologías de la información y las comunicaciones en general y las comunicaciones móviles en particular tienen una incidencia decisiva en el crecimiento económico, la competitividad y la mejora de la productividad. El terminal móvil ha llegado a constituir hoy en día una parte esencial en la esfera de objetos personales. En este contexto, la industria de las comunicaciones móviles ha venido aportando soluciones al mercado, en la forma de sucesivas generaciones de sistemas. La globalización de los mercados y la búsqueda de economías de escala son algunos de los principales argumentos (esgrimidos ya en la concepción de la segunda generación de comunicaciones) que justifican el interés y el desarrollo de sistemas estándares, resultado del consenso entre los diferentes agentes implicados. Así, los diferentes organismos y foros de estandarización adquieren una relevancia muy significativa en el marco general del negocio de las comunicaciones móviles.

Los procesos de estandarización son costosos en tiempo y esfuerzo. Típicamente se inician con una primera fase en la que se establecen los requisitos que debe satisfacer el diseño del sistema que se pretende estandarizar. Tras ello, se decide la arquitectura del sistema, con sus principales bloques y correspondientes interfaces. Sobre ello, se procede a la especificación detallada, así como el test y la verificación que pueden iniciarse cuando las especificaciones alcanzan ya un alto nivel de estabilidad. El proceso es iterativo. Por ejemplo, pueden añadirse, modificarse o eliminarse requisitos a la vista de las soluciones técnicas que se vayan formulando. Similarmente, pueden modificarse las soluciones técnicas si se evidencian dificultades en la verificación práctica.

1.3.1 LTE

En el caso de LTE, las especificaciones emanan del 3GPP (3rd Generation Partnership Project), que nació en 1998 con el objetivo de especificar 3G (UTRA-FDD y UTRA-TDD). También se encarga de mantener y desarrollar las especificaciones de GERAN (GSM EDGE RAN). La red de acceso radio se especifica en el marco del TSG RAN, que se organiza en cinco grupos de trabajo: WG1 (capa física), WG2 (capas 2 y 3), WG3 (interfaces fijos de la red de acceso), WG4 (aspectos de RF y RRM) y WG5 (conformidad de terminales). Los documentos del 3GPP se estructuran en *Releases*, cada una de ellas caracterizada por la incorporación de un conjunto de funcionalidades destacadas en relación a la versión anterior. Así, la que se llamó R99 (por el hecho de que se congeló en diciembre de 1999) supuso el primer conjunto de especificaciones UMTS. Seguidamente, tras la llamada R4, se completó en marzo de 2002 la R5 que incluye por ejemplo HSDPA. Tres años después se incorpora HSUPA así como MBMS en R6. En la R7 (septiembre de 2007) se incluye HSPA+, mientras que LTE/SAE se asocian ya a R8 y posteriores.

Puede decirse que el primer paso hacia LTE se llevó a cabo en noviembre de 2004, cuando 3GPP TSG RAN organizó un *Workshop* sobre "RAN Evolution" en Toronto (Canadá), en el que se presentaron unas 40 contribuciones con ideas, propuestas, etc. En el propio *Workshop*

se identificaron una serie de requisitos de alto nivel, como un coste por bit reducido, mejora en la provisión de servicios, flexibilidad en el uso de las bandas frecuenciales, arquitectura simplificada con interfaces abiertos, consumo de potencia en el terminal razonable, etc. También se puso de manifiesto que el esfuerzo de estandarización que esta evolución, bautizada como E-UTRAN (Evolved UTRAN), llevaría asociado sólo resultaría justificable si las mejoras fueran significativas.

En diciembre de 2004 se creó el *Study Item* "Evolved UTRA and UTRAN" para la evolución hacia una tecnología de acceso de elevada velocidad de transmisión, baja latencia y optimizada para la transmisión de paquetes, de modo que con ello quedase asegurada la competitividad de las soluciones 3GPP en un horizonte temporal largo. En particular, algunos de los objetivos de E-UTRA y E-UTRAN son [34]:

- Velocidades de transmisión de pico de 100 Mbps en *downlink* y 50 Mbps en *uplink*, mejorando la velocidad de transmisión obtenible en el extremo de la célula.
- Mejora de la eficiencia espectral en un factor 2-4 con respecto a la Release 6.
- Latencia del plano de usuario en la red de acceso radio inferior a 10 ms.
- Ancho de banda escalable.
- Interoperabilidad con sistemas 3G y sistemas no 3GPP.

Este primer periodo de trabajo concluyó con la aprobación del TR 25.912 [35] en la reunión TSG-RAN #32 en junio de 2006. En TR 25.912 se recomendaba la creación de un *Work Item* sobre E-UTRA y E-UTRAN tomando como punto de partida el concepto de sistema reflejado en dicho documento, ya que los estudios realizados validaban su viabilidad. Tras considerarse y discutirse múltiples propuestas sobre la capa física, protocolos radio, arquitectura de red, aspectos de RF, consideraciones de complejidad, etc., algunas de las características principales incorporadas al concepto del TR 25.912 fueron:

- Esquema de acceso radio OFDMA en el *downlink* y SC-FDMA en el *uplink*.
- Soporte de *packet scheduling* en el dominio temporal y frecuencial.
- Simplificaciones en la MAC y en el modelo de estados RRC, así como reducción del número de canales de transporte (no hay canales dedicados).
- Funcionalidades de *packet scheduling*, ARQ e HARQ terminadas en eNodeB.
- Simplificación de la arquitectura E-UTRAN y descentralización de la misma.

A partir de este *Feasibility Study* se pasó al desarrollo de tareas de especificación (*Work Items*). Para cada WI se establece un título y ámbito técnico, se define su resultado esperado (*Technical Specification o Technical Report*) e impacto sobre otros WIs así como el calendario de tareas. La primera versión completa de las especificaciones LTE se aprobó en diciembre de 2007. Durante 2008 el trabajo del 3GPP se centró en la finalización de *Release 8*, aunque también se han ido desarrollando con intensidad aspectos de *Release 9* y *Release 10*, como

se comentará en el apartado 1.3.2. Las especificaciones de *Release 8* quedaron congeladas en diciembre de 2008, lo que significa que no se pueden añadir nuevas funcionalidades a esta versión, si bien se continúa el trabajo para completar los contenidos acordados.

Cabe destacar que las capacidades y prestaciones de E-UTRA y E-UTRAN establecidas en [34] corresponden a los objetivos fijados en la fase inicial de desarrollo de LTE, de manera que las prestaciones finalmente alcanzadas en muchos casos pueden superar los objetivos iniciales. Por ejemplo, la velocidad de pico teórica alcanzable en el *downlink* para 2×20 MHz (FDD), 64-QAM y 4×4 MIMO resulta de 326 Mbits/s [36]. No obstante, si bien por la propia necesidad de sintetizar las capacidades de un sistema se tiende a proporcionar la velocidad de transmisión de pico, en términos de evaluación de prestaciones las principales métricas a considerar deben ser de otro tipo (por ejemplo, el percentil 95 de la distribución bits/s/MHz/célula). Así mismo, siendo una buena referencia la cota teórica a nivel de enlace, es fundamental evaluar las prestaciones a nivel de sistema en entornos realistas (a través de simulaciones en una primera instancia y medidas de campo en una segunda). Véase [37] como ejemplo.

En el contexto de la ITU, UTRA y E-UTRA (LTE) son el pariente europeo de la familia IMT-2000, ya que en realidad IMT-2000 no es una tecnología de acceso radio en sí misma, sino una familia de tecnologías que cumplen los requisitos establecidos por la ITU para IMT-2000 y que son aprobadas por la propia ITU. La principal recomendación IMT-2000 es ITU-R M. 1457[38], en la que se incluye una descripción de cada uno de los miembros de la familia IMT-2000 acompañada de una lista de referencias a las especificaciones detalladas. ITU-R WP5D se encarga de ir revisando dicha recomendación, dado el continuo desarrollo de los diferentes interfaces radio IMT-2000.

Cabe mencionar que en ITU-R M. 1457-7 (Revisión 7 del documento), aprobado en octubre de 2007, se incluyó también el estándar IEEE 802.16 (WiMAX), como un nuevo miembro de IMT-2000. Dentro también de las familias IMT-2000 se encuentra CDMA-2000 y UMB (Ultra Mobile Broadband), que constituyen el equivalente a UMTS y LTE en el marco de 3GPP2, y que en el recorrido desde 2G (con IS-95) hasta la evolución de 3G ha supuesto el principal polo de competencia a las tecnologías 3GPP. No obstante, el desarrollo de UMB ha quedado paralizado desde finales de 2008, cuando algunos de los principales motores de las comunicaciones móviles (Qualcomm, Verizon, etc.) se inclinaron por LTE.

1.3.2 LTE Advanced

Al tiempo que las especificaciones de UMTS R99 se estaban completando, la ITU inició las primeras consideraciones en la concepción de soluciones más allá de IMT-2000, conocido actualmente como IMT-Advanced y referido como 4G. La Resolución 228 (WRC-2000 y revisada en WRC-03) invitó a la ITU-R a estudiar en detalle los aspectos técnicos y operativos relativos al futuro desarrollo de IMT-Advanced.

El marco general y los objetivos están definidos en la Recomendación ITU-R M.1645[39].

Por su parte, la Resolución ITU-R 57 [40] establece los principios y criterios esenciales para el desarrollo de Recomendaciones para la interfaz radio IMT-Advanced. La ITU-R Circular Letter 5/LCCE/2 de marzo de 2008 [41] invita a la presentación de propuestas candidatas para la interfaz radio *IMT-Advanced*. La Revision 1 del documento IMT-ADV/2-E de agosto de 2008 [42] establece el calendario de desarrollo de IMT-Advanced. El 7 de octubre de 2009 el 3GPP presentó la propuesta "LTE Release 10 & beyond (*LTE-Advanced*)", completando así el *Step 3* de dicho proceso (recepción de candidatos). El documento de referencia para el 3GPP en este punto lo constituye la TR 36.912 v9.0.0 [43].

Previendo el inicio del proceso IMT-Advanced en el seno de ITU, el 3GPP inició en marzo de 2008 el Study Item LTE-Advanced [44]. El calendario del 3GPP está alineado con el calendario de la ITU, de manera que se espera completar las especificaciones de la Release 10 en la primera parte de 2011. Sin embargo, la visión dentro del 3GPP es que LTE-Advanced no debe limitarse a cumplir los requerimientos de IMT-Advanced, sino que debe perseguir unos objetivos mucho más ambiciosos.

LTE se concibe como el punto de inicio para una transición suave hacia el acceso radio 4G (esto es, IMT-Advanced) o, en otras palabras, LTE-Advanced es la evolución de LTE. En este sentido, LTE-Advanced debe asegurar toda una serie de requisitos en relación a la compatibilidad hacia atrás con LTE *Release 8*. En cuanto a compatibilidad espectral, LTE-Advanced debería poderse desplegar en bandas ocupadas por LTE. Así mismo, el equipamiento LTE debería poder incorporar las funcionalidades LTE-Advanced con una complejidad y coste razonablemente bajos.

Los requerimientos de LTE-Advanced establecidos en TR 36.913 [45] distinguen diferentes categorías: generales, capacidades (velocidad de transmisión de pico, latencias), prestaciones del sistema (eficiencia espectral, *throughput* en el extremo de la célula, movilidad, cobertura, etc.), despliegue (espectro, coexistencia e interoperación con *legacy* RATs, etc.), arquitectura E-UTRAN y migración, complejidad, coste, etc.

Para poder satisfacer los requerimientos establecidos (por ejemplo, soporte de velocidades de pico de hasta 1 Gbit/s en *downlink* y 500 Mbit/s en *uplink*), son necesarias una serie de mejoras técnicas con respecto a LTE (*Release 8*). Algunas de las principales componentes técnicas de LTE-Advanced son [43]:

- Agregación de banda hasta 100 MHz, por ejemplo a partir de agregar múltiples componentes de 20 MHz para poder alcanzar un ancho de banda de 100 MHz y así proporcionar las velocidades de transmisión más elevadas previstas en los requerimientos.
- Extensión de soluciones multi-antena, con hasta 8 niveles en el *downlink* y 4 niveles en el *uplink*, para así incrementar las velocidades de transmisión alcanzables sobre el enlace.
- *Coordinated multipoint transmission and reception* (CoMP), que permite mejorar las prestaciones observables en el extremo de la célula a través de efectuar la transmi-

sión/recepción desde distintas células. CoMP es un término relativamente general, que incluye diferentes tipos de coordinación (*packet scheduling*, *beam-forming*, etc.) entre transceptores separados geográficamente.

- Repetidores, como mecanismo para mejorar la cobertura y reducir el coste de despliegue.

1.4 Referencias bibliográficas

- [1] ABI Research, "Femtocell Access Points: Fixed-Mobile Convergence for Residential, SMB, and Enterprise Markets", agosto de 2006.
- [2] C. E. Shannon, "A Mathematical Theory of Communications", Bell System Technical Journal, Vol. 27, julio y octubre de 1948.
- [3] O. Sallent, J. L. Valenzuela, R. Agustí, *Principios de Comunicaciones Móviles*, Ediciones UPC, 2003.
- [4] L. J. Greenstein, et al., "Microcells in personal communications systems", IEEE Communications Magazine, diciembre de 1992.
- [5] IEEE Communications Magazine, Feature Topic: Femtocells, septiembre de 2009.
- [6] J. Pérez-Romero, O. Sallent, R. Agustí, M. A. Díaz-Guerra, Radio resource management strategies in UMTS, John Wiley & Sons, 2005.
- [7] J. Pérez-Romero, O. Sallent, R. Agustí, "On the optimum traffic allocation in heterogeneous CDMA/TDMA networks", IEEE Transactions on Wireless Communications, Vol. 6, nº 9, septiembre de 2007.
- [8] NGMN Alliance Deliverable, NGMN Use Cases related to Self Organising Network, Overall Description, May 2007.
- [9] 3GPP TR 32.816, Study on management of Evolved Universal Terrestrial Radio Access Network (E-UTRAN) and Evolved Packet Core (EPC).
- [10] 3GPP TR 36.902, Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Self-configuring and self-optimizing network use cases and solutions.
- [11] Directiva 87/372/CEE del Consejo de 25 de junio de 1987 relativa a las bandas de frecuencia a reservar para la introducción coordinada de comunicaciones móviles terrestres digitales celulares públicas paneuropeas en la Comunidad.
- [12] Commission Decision 2009/766/EC de 16 de octubre de 2009 "on the harmonisation of the 900 MHz and 1 800 MHz frequency bands for terrestrial systems capable of providing pan-European electronic communications services in the Community".
- [13] Viviane Reding (Member of the European Commission responsible for Information Society and Media), Speech/08/476, "The digital dividend: A unique opportunity for Europe's wireless economy", octubre de 2008.

- [14] Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, COM(2007) 700, "Aprovechar plenamente las ventajas del dividendo digital en Europa: un planteamiento común del uso del espectro liberado por la conversión al sistema digital", noviembre de 2007.
- [15] Commission study "A European approach to the digital dividend", realizado por Analysys Mason, DotEcon and Hogan&Hartson.
- [16] European Commission, Information Society and Media Directorate-General, Consultation document, "Transforming the digital dividend opportunity into social benefits and economic growth in Europe", julio de 2009.
- [17] Commission Recommendation 2009/848/EC, "Facilitating the release of the digital dividend in the European Union", 28 de octubre de 2009.
- [18] Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, COM(2009) 586 2, "Transforming the digital dividend into social benefits and economic growth", 28 de octubre de 2009.
- [19] Ofcom, Digital Dividend Review, A statement on our approach to awarding the digital dividend, diciembre de 2007.
- [20] Federal Communications Commission (FCC), "Spectrum Policy Task Force," ET Docket no. 02-135, november 15, 2002.
- [21] P. Kolodzy, "Spectrum Policy Task Force: Findings and Recommendations", International Symposium on Advanced Radio Technologies (ISART), March 2003.
- [22] J. Mitola et al., "Cognitive Radios: Making Software Radios more Personal," IEEE Personal Communications, Vol. 6, n°. 4, August 1999.
- [23] Federal Communications Commission (FCC), "Notice of Proposed Rule Making," ET Docket no. 04-113, may 25, 2004.
- [24] Ofcom, Digital dividend: Cognitive Access, Consultation on licence-exempting cognitive devices using interleaved spectrum, febrero de 2009.
- [25] Ofcom, Digital dividend: cognitive access, Statement on licence-exempting cognitive devices using interleaved spectrum, julio de 2009.
- [26] K.S. Gilhousen et al., "On the Capacity of a Cellular CDMA System", IEEE Transactions on Vehicular Technology, Vol. 40, No. 2, mayo 1991.
- [27] T. Ojanpera, et al. "Analysis of CDMA and TDMA for 3rd generation mobile radio systems", IEEE 47th Vehicular Technology Conference, Phoenix (EEUU), mayo de 1997.
- [28] T. Halonen, et al., GSM, GPRS and EDGE Performance – Evolution towards 3G/UMTS, John Wiley & Sons, 2002.
- [29] H. Holma, A. Toskala, HSDPA/HSUPA for UMTS – High Speed Radio Access for Mobile

Communications, John Wiley & Sons, 2006.

- [30] C. Berrou, A. Glavieux, P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes", IEEE International Conference on Communications, 1993 (ICC'93).
- [31] O. Y. Takeshita, "On Maximum Contention-Free Interleavers and Permutation Polynomials Over Integer Rings", IEEE Transactions on Information Theory, Vol.52, n° 3, march 2006.
- [32] M. Mouly, M. Pautet, The GSM System for Mobile Communication, 1992.
- [33] J. Foschini, M. J. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas", Wireless Personal Communications 6:1998.
- [34] 3GPP TR 25.913, Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN).
- [35] 3GPP TR 25.912, Feasibility study for evolved Universal Terrestrial Radio Access (UTRA) and Universal Terrestrial Radio Access Network (UTRAN).
- [36] 3GPP R1-072578, "Summary of Downlink Performance Evaluation", Ericsson, mayo de 2007.
- [37] J. Karlsson, M. Riback, "Initial field performance measurements of LTE", Ericsson Review, 3 2008.
- [38] Recommendation ITU-R M. 1457. Detailed specifications of the radio interfaces of International Mobile Telecommunications-2000 (IMT-2000).
- [30] Recommendation ITU-R M.1645 (2003), Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000.
- [40] Resolution ITU-R 57 (2007), "Principles for the process of development of IMT-Advanced".
- [41] ITU-R Circular Letter 5/LCCE/2, "Invitation for submission of proposals for candidate radio interface technologies for the terrestrial components of the radio interface(s) for IMT-Advanced and invitation to participate in their subsequent evaluation".
- [42] ITU-R, Revision 1 to Document IMT-ADV/2-E, Submission and evaluation process and consensus building .
- [43] 3GPP TR 36.912 v9.0.0, Feasibility study for Further Advancements for EUTRA (LTE-Advanced) (Release 9)].
- [44] 3GPP TD RP-080137: "Proposed SID on LTE-Advanced", NTT DoCoMo.
- [45] 3GPP TR 36.913, "Requirements for further advancements for E-UTRA (LTE-Advanced)".



II

ARQUITECTURA DEL SISTEMA LTE

2. Arquitectura del sistema LTE

2.1 Introducción

En este capítulo se describe la arquitectura de una red de comunicaciones móviles basada en las especificaciones del sistema LTE. Para ello, a partir de un análisis inicial de la arquitectura genérica adoptada en los sistemas celulares 2G/3G actuales, se identifican las piezas fundamentales que componen la arquitectura de red de toda la familia de sistemas especificada por 3GPP (GSM, UMTS, LTE). Esta identificación permite acotar de forma clara cuáles son, y a qué criterios básicos de diseño obedecen, los nuevos componentes introducidos por el sistema LTE respecto a las redes GSM y UMTS.

Una vez identificados los componentes de más alto nivel que forman parte del sistema LTE, en posteriores apartados se realiza una descripción detallada de cada uno de ellos en base a las entidades de red (e.g., estaciones base, pasarela de red, etc.) e interfaces asociadas en que se estructuran internamente. De cada una de las entidades de red se indican sus funciones más relevantes y se proporcionan las referencias necesarias hacia otros apartados del libro donde se abordan con más detalle algunas de sus funcionalidades. Respecto a las interfaces entre las entidades de red, conjuntamente con la descripción de su funcionalidad, se describen las torres de protocolos que sustentan las interfaces y los principios básicos de los protocolos utilizados.

Aclaraciones sobre nomenclatura

El término LTE se acuñó inicialmente en 3GPP para denominar una línea de trabajo interna cuyo objeto de estudio era la evolución de la red de acceso de UMTS, denominada como UTRAN. Formalmente, la nueva red de acceso recibe el nombre de E-UTRAN (Evolved UTRAN) aunque muchas veces se utiliza también el término LTE en las especificaciones como sinónimo de E-UTRAN. Asimismo, en lo concerniente a la red troncal, 3GPP utilizó el término SAE (System Architecture Evolution) para referirse a las actividades de estudio relacionadas con la especificación de una red troncal evolucionada de conmutación de paquetes. Formalmente, dicha red troncal se denomina EPC (Evolved Packet Core) o también *Evolved 3GPP Packet Switched Domain*, y de la misma forma que pasa con la red de acceso, es común encontrar el término de SAE como sinónimo de EPC. La combinación de la red de acceso E-UTRAN y la red troncal EPC es lo que constituye la nueva red UMTS evolucionada y recibe el nombre formal de EPS (*Evolved Packet System*). La primera especificación del sistema EPS ha sido incluida en la Release 8 de las especificaciones del 3GPP.

Clarificadas las formalidades de los términos LTE, SAE, E-UTRAN, EPC y EPS en el contexto del trabajo y especificaciones del 3GPP, es importante tener en cuenta que ETSI ha registrado "LTE", y su logotipo asociado, como marca comercial para hacer referencia de forma clara a la nueva red UMTS evolucionada. Por ello, es común encontrar la utilización del término LTE como sinónimo de EPS. En este libro se ha adoptado también este criterio.

2.2 Arquitectura genérica de los sistemas celulares

En la Figura 2.1 se ilustra una arquitectura simplificada de un sistema de comunicaciones móviles celular. Esta arquitectura representa un modelo de la red a muy alto nivel donde se identifican tres componentes básicos:

- Equipo de usuario, dispositivo que permite al usuario acceder a los servicios de la red. El equipo de usuario puede incluir una tarjeta inteligente (Universal Integrated Circuit Card, UICC) que contenga la información necesaria para permitir la conexión a la red y la utilización de sus servicios (e.g., identificador único del usuario en el sistema de comunicaciones). El equipo de usuario se conecta a la red de acceso a través de una interfaz radio.
- Red de acceso, parte del sistema responsable de sustentar la transmisión radio con los equipos de usuario de cara a proporcionar la conectividad necesaria entre éstos y los equipos de la red troncal. Los servicios de transmisión ofrecidos por la red de acceso para transportar la información de los equipos de usuario (tanto información de datos como señalización) hacia/desde la red troncal son servicios portadores, es decir, servicios cuya finalidad última es la provisión de una cierta capacidad de transmisión. La red de acceso es la responsable de gestionar el uso de los recursos radio disponibles para la provisión de servicios portadores de forma eficiente. La activación de los recursos de transmisión en la red de acceso se controla generalmente desde la red troncal. La red de acceso está formada por estaciones base y, en los sistemas móviles actuales 2G y 3G, también por equipos controladores de las estaciones base.
- Red troncal, parte del sistema encargado de aspectos tales como control de acceso a la red celular (e.g., autenticación de los usuarios del sistema), gestión de la movilidad de los usuarios, gestión de las sesiones de datos o circuitos que transportan la información de los usuarios, mecanismos de interconexión con otras redes, etc. También pueden formar parte de la red troncal las funciones asociadas con el control de los servicios finales¹ ofrecidos a los usuarios (e.g., control y señalización asociada al servicio de telefonía). La red troncal está formada por equipos que albergan funciones de conmutación de circuitos, encaminamiento de paquetes (*routing*), bases de datos, etc.

Esta arquitectura genérica ha sido adoptada en las diferentes familias de sistemas celulares 2G y 3G, y también se mantiene en el sistema LTE. La separación entre la red de acceso y red troncal confiere un importante grado de flexibilidad al sistema de cara a soportar un proceso evolutivo en el que se puedan ir mejorando, agregando o sustituyendo las diferentes partes de la red con la mínima afectación posible al resto de la misma.

¹ Los servicios *finales* son los diferentes servicios de comunicación (e.g., telefonía, videoconferencia, etc.) que se ofrecen a través del sistema. El calificativo *finales* se utiliza en este libro para enfatizar la diferencia entre estos servicios y los servicios *portadores* . Un servicio portador básicamente se concibe como un servicio de transporte de información entre dos puntos de la red. Así, el transporte de la información generada por los servicios finales se realiza a través de los servicios portadores que ofrece la red.

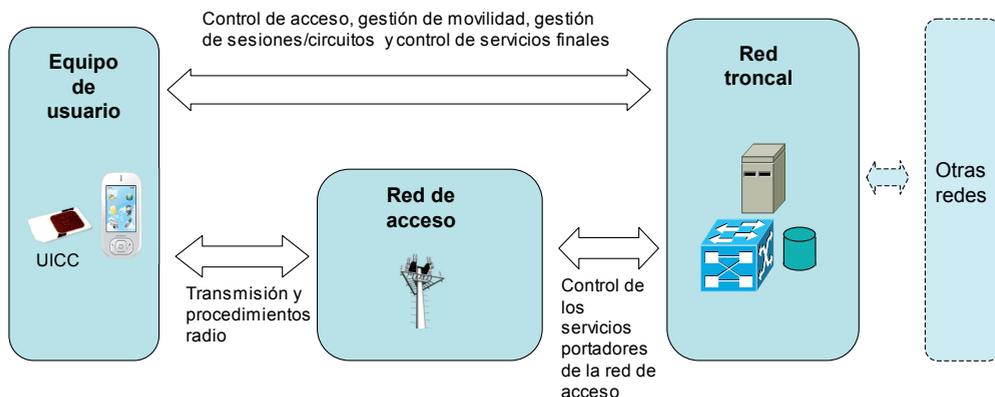


Figura 2.1 Arquitectura genérica de un sistema celular

En la Figura 2.2 se muestran dos escenarios representativos de la provisión de servicios de comunicación a través de redes de comunicaciones móviles celulares. En el escenario (a) se ilustra una red celular que sustenta servicios de comunicación (e.g., llamadas de voz) entre los equipos de usuario conectados a ella. En el escenario (b) se muestra la provisión de los servicios de comunicación entre equipos de usuario operando en redes celulares diferentes interconectadas entre sí mediante redes de tránsito. En la misma línea, en la Figura 2.3 se representa un escenario ilustrativo de la provisión de servicios entre equipos conectados a redes celulares y equipos localizados en otras redes (e.g., red telefónica, Internet, etc.).

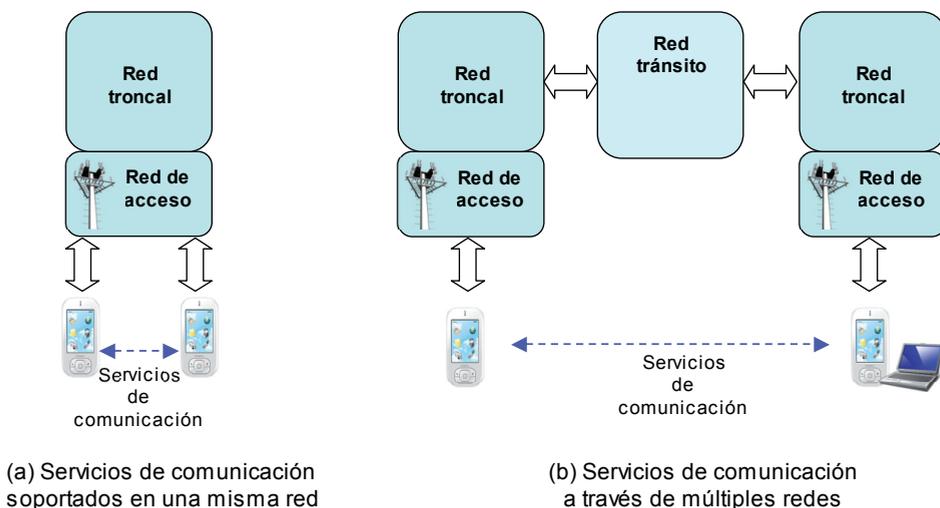


Figura 2.2 Escenarios de provisión de servicios de comunicación en redes celulares

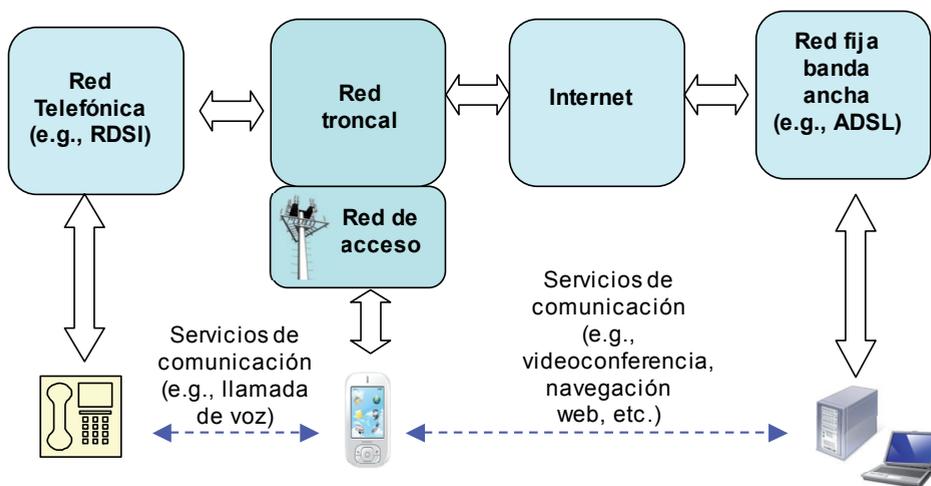


Figura 2.3 Escenario de provisión de servicios de comunicación entre redes celulares y otras redes

2.3 Arquitectura general de los sistemas 3GPP

Las arquitecturas de red contempladas en la familia de sistemas especificados por 3GPP [1] se adaptan a la arquitectura genérica descrita en el apartado anterior. Así pues, tal como se representa en la Figura 2.4, los sistemas 3GPP abarcan la especificación del equipo de usuario (*User Equipment*, UE) y de una infraestructura de red que se divide de forma lógica en una infraestructura de red troncal (*Core Network*, CN) y una de red de acceso (*Access Network*, AN).

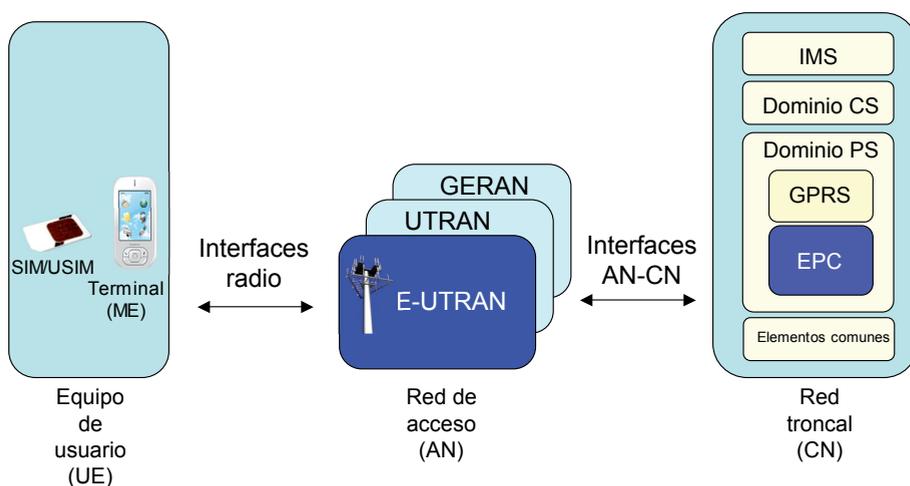


Figura 2.4 Arquitectura de alto nivel de los sistemas 3GPP (GSM, UMTS y LTE)

El equipo de usuario en 3GPP se compone de dos elementos básicos: el propio dispositivo móvil o terminal (denominado como *Mobile Equipment*, ME, en las especificaciones) y una tarjeta UICC. La tarjeta UICC, también denominada SIM (*Subscriber Identity Module*) en sistemas GSM y USIM (*Universal SIM*) en UMTS y LTE, es la encargada de almacenar la información y sustentar los procedimientos que tienen que ver con la suscripción del usuario a los servicios proporcionados por la red. Mediante esta separación entre terminal y tarjeta se permite que un usuario (identificado a través de la SIM/USIM) pueda utilizar diferentes terminales para acceder a la red. En el apartado 2.8 se cubren con más detalle las características de los equipos de usuario 3GPP.

Respecto a la red de acceso, 3GPP ha especificado tres tipos de redes de acceso diferentes: GERAN (*GSM/EDGE Radio Access Network*), UTRAN (*UMTS Terrestrial Radio Access Network*) y E-UTRAN (*Evolved UTRAN*). Las redes de acceso GERAN y UTRAN forman parte del sistema 3G UMTS mientras que E-UTRAN es la nueva red de acceso del sistema LTE. Cada red de acceso define su propia interfaz radio para la comunicación con los equipos de usuario: GERAN, también denominada de forma habitual simplemente como GSM, utiliza un acceso basado en TDMA, la tecnología utilizada en UTRAN es WCDMA y, E-UTRAN ha apostado por la tecnología OFDMA. Asimismo, la interconexión de las redes de acceso a la red troncal se realiza mediante interfaces AN-CN específicas a cada una de ellas.

Respecto a la red troncal, ésta se divide de forma lógica en un dominio de circuitos (*Circuit Switched, CS, Domain*), un dominio de paquetes (*Packet Switched, PS, Domain*) y el subsistema IP Multimedia (*IP Multimedia Subsystem, IMS*). En adelante, nos referiremos a estos tres componentes como dominio CS, dominio PS y subsistema IMS, respectivamente.

El dominio CS alberga a todas las entidades de la red troncal que participan en la provisión de servicios de telecomunicación basados en conmutación de circuitos, es decir, servicios a los que se les asignan recursos de forma dedicada (circuitos) en el momento de establecimiento de la conexión, manteniéndose éstos hasta la finalización del servicio (e.g., servicios de voz y videoconferencia en redes UMTS). El dominio de circuitos de la red troncal es accesible a través de las redes de acceso UTRAN y GERAN. En cambio, el diseño de E-UTRAN no contempla el acceso al dominio CS ya que todos los servicios se proporcionan a través del dominio PS.

El dominio PS incluye a las entidades de la red troncal que proporcionan servicios de telecomunicación basados en conmutación de paquetes: la información de usuario se estructura en paquetes de datos que se encaminan y transmiten por los diferentes elementos y enlaces de la red. En particular, el dominio PS proporciona un servicio de conectividad a redes de paquetes (e.g., redes IP y X.25). Existen dos implementaciones diferentes del dominio PS: GPRS y EPC. GPRS es la implementación del dominio PS que se desarrolló inicialmente en el contexto de redes GSM y que actualmente también forma parte del sistema UMTS. Los servicios de conectividad por paquetes de GPRS son accesibles tanto a través de UTRAN como de GERAN. Por otro lado, EPC es la nueva especificación del dominio PS desarrollada en el contexto del sistema LTE. Tal como se verá más adelante, EPC es una implementación evolucionada de

GPRS que ha sido optimizada para proporcionar un servicio de conectividad IP a los equipos de usuario a través de E-UTRAN. El dominio EPC también ha sido concebido para soportar el acceso al servicio de conectividad IP desde las otras redes de acceso 3GPP (UTRAN y GERAN) así como desde redes no 3GPP (e.g., redes CDMA2000, Mobile WiMAX, etc.).

En la red troncal existen también elementos que soportan funciones asociadas a los diferentes dominios. Un ejemplo claro es la base de datos que contiene la información de los usuarios del sistema (HSS, *Home Subscriber Server*) y sobre la que se sustenta la operación de los dominios CS y PS, así como del subsistema IMS comentado a continuación.

También como parte de la red troncal, el subsistema IMS comprende los elementos de ésta relacionados con la provisión de servicios IP multimedia basados en el protocolo SIP (*Session Initiation Protocol*) de IETF (*Internet Engineering Task Force*²). El subsistema IMS es responsable de la señalización asociada a los servicios multimedia y utiliza como mecanismo de transporte los servicios de transferencia de datos proporcionados por el dominio PS (i.e., el equipo de usuario y los equipos del subsistema IMS o redes externas se comunican entre sí a través del servicio de conectividad IP ofrecido por el dominio PS). En este sentido, el subsistema IMS constituye el plano de control de dichos servicios quedando claramente separadas las funciones asociadas al transporte de la información (funciones asociadas al dominio PS) y las funciones propias de la capa de control de los servicios (señalización a nivel de aplicación). Esta separación entre las funciones de transporte y las de control de los servicios adoptada en LTE está en consonancia con los trabajos desarrollados en otros organismos de estandarización como ETSI (TISPAN [2]) y ITU-T (NGN-GSI [3]) respecto a la especificación de las arquitecturas de los futuros sistemas de telecomunicación basados íntegramente en redes de conmutación de paquetes y que responden a la denominación de *Next Generation Networks* (NGN).

La arquitectura completa de la familia de estándares 3GPP se proporciona en la especificación TS 3GPP 23.002 [1].

2.4 Arquitectura del sistema LTE

Atendiendo a la arquitectura general de los sistemas 3GPP, en la Figura 2.5 se ilustra de forma simplificada la arquitectura completa del sistema LTE, denominado formalmente en las especificaciones como *Evolved Packet System* (EPS). Los componentes fundamentales del sistema LTE son, por un lado, la nueva red de acceso E-UTRAN y el nuevo dominio de paquetes EPC de la red troncal (denominado en adelante simplemente como red troncal EPC), y por otro, la evolución del subsistema IMS concebido inicialmente en el contexto de los sistemas UMTS. Los diferentes componentes han sido diseñados para soportar todo tipo de servicios de telecomunicación mediante mecanismos de conmutación de paquetes, por lo que no resulta necesario disponer de un componente adicional para la provisión de servicios

² La misión de IETF (<http://www.ietf.org/>) es el desarrollo de protocolos y otras especificaciones técnicas que permitan la evolución de Internet y faciliten su gestión.

en modo circuito (en el sistema LTE los servicios con restricciones de tiempo real se soportan también mediante conmutación de paquetes). En este sentido, EPC constituye una versión evolucionada del sistema GPRS.

La red de acceso E-UTRAN y la red troncal EPC proporcionan de forma conjunta servicios de transferencia de paquetes IP entre los equipos de usuario y redes de paquetes externas tales como plataformas IMS y/o otras redes de telecomunicaciones como Internet. Las prestaciones de calidad de servicio (e.g., tasa de datos en bits/s, comportamientos en términos de retardos y pérdidas) de un servicio de transferencia de paquetes IP puede configurarse en base a las necesidades de los servicios finales que lo utilicen, cuyo establecimiento (señalización) se lleva a cabo a través de plataformas de servicios externas (e.g., IMS) y de forma transparente a la red troncal EPC. Formalmente, el servicio de transferencia de paquetes IP ofrecido por la red LTE entre el equipo de usuario y una red externa se denomina servicio portador EPS (*EPS Bearer Service*). Asimismo, la parte del servicio de transferencia de paquetes que proporciona la red de acceso E-UTRAN se denomina E-UTRAN *Radio Access Bearer* (E-RAB). El modelo de conectividad completo del sistema LTE basado en servicios portadores se aborda más detalladamente en el apartado 3.2 del Capítulo 3.

En la Figura 2.5 se muestran las principales interfaces de E-UTRAN y EPC. Tal como se verá en el apartado 2.5, la interfaz entre E-UTRAN y EPC se denomina S1 y proporciona a la EPC los mecanismos necesarios para gestionar el acceso de los terminales móviles a través de E-UTRAN. La interfaz radio entre los equipos de usuario y E-UTRAN se denomina E-UTRAN U_{μ} . Por otro lado, las plataformas de servicios como IMS y la conexión a redes de paquetes externas IP se lleva a cabo mediante la interfaz SGi de la EPC. La interfaz SGi es análoga a la interfaz Gi definida en las redes GPRS/UMTS y constituye el punto de entrada/salida al servicio de conectividad IP proporcionado por la red LTE (los terminales conectados a la red LTE son “visibles” a las redes externas a través de esta interfaz mediante su dirección IP). Los mecanismos de control de los servicios de transporte ofrecidos por EPC se sustentan en información proporcionada por otros elementos de la red troncal que no son exclusivos del sistema LTE sino que pueden dar soporte también a otros dominios de los sistemas 3GPP. En la Figura 2.5 se mencionan algunos de estos elementos comunes entre los que, a modo de ejemplo, se encuentra la base de datos del sistema con la información de suscriptores de sus usuarios (HSS). La funcionalidad de todos estos elementos comunes se describirá más detalladamente en el apartado 2.6 dedicada a cubrir la red troncal EPC.

Otra característica fundamental del sistema LTE es que contempla también el acceso a sus servicios a través de UTRAN y GERAN así como mediante la utilización de otras redes de acceso que no pertenecen a la familia 3GPP (e.g., CDMA2000, Mobile WiMAX, redes 802.11, etc.). La interconexión de las redes de acceso alternativas, tanto 3GPP como no, se soporta a través de un conjunto de interfaces de la EPC. En el Capítulo 7 se analizan más detalladamente todos los aspectos relacionados con el *interworking* de diferentes redes de acceso radio.

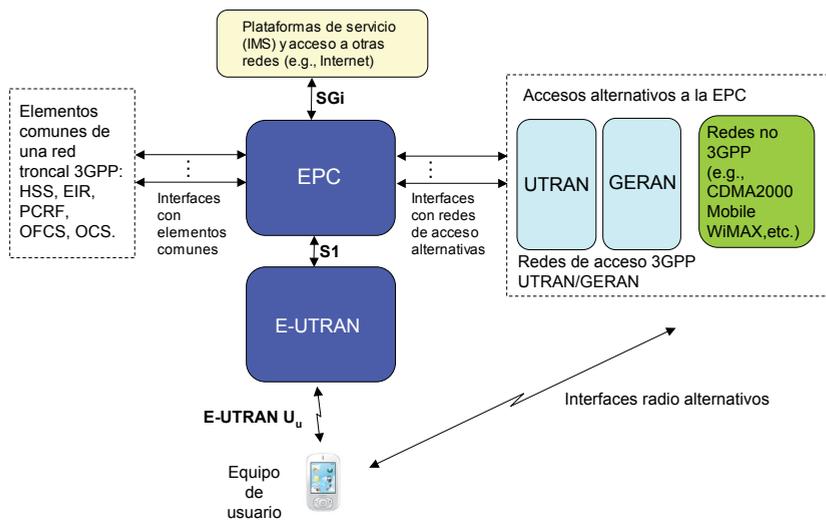


Figura 2.5 Arquitectura del sistema LTE

Finalmente, aunque no quede reflejado en la Figura 2.5, es importante destacar que la interconexión de los diferentes equipos físicos donde se ubicarían las funciones tanto de la red troncal EPC como de la red de acceso E-UTRAN, se realiza mediante tecnologías de red basadas en IP. De esta forma, la red física que se utiliza para interconectar los diferentes equipos de una red LTE, y que se denomina comúnmente como red de transporte, es una red IP convencional. Por tanto, la infraestructura de una red LTE, además de los equipos propios que implementan las funciones del estándar 3GPP, también integra otros elementos de red propios de las redes IP tales como *routers*, servidores DHCP (*Dynamic Host Configuration Protocol*) para la configuración automática de las direcciones IP de los equipos de la red LTE y servidores DNS (*Domain Name Server*) para asociar los nombres de los equipos con sus direcciones IP.

2.5 Red de acceso evolucionada: E-UTRAN

2.5.1 Arquitectura de E-UTRAN

La arquitectura de la red de acceso se compone de una única entidad de red³ denominada *evolved NodeB* (eNB) que constituye la estación base de E-UTRAN. Así pues, la estación base E-UTRAN integra toda la funcionalidad de la red de acceso, a diferencia de las redes de acceso de GSM y UMTS compuestas por estaciones base (BTS, NodoB) y equipos controladores (BSC y RNC). La descripción de la arquitectura de E-UTRAN se detalla en las especificaciones del 3GPP TS 36.300 [4] y TS 36.401 [5].

³ Formalmente, una entidad de red en 3GPP representa una entidad lógica que cubre una funcionalidad perfectamente delimitada. Por tanto, una entidad de red es una entidad funcional.

Tal y como se ilustra en la Figura 2.6, una red de acceso E-UTRAN está formada por eNBs que proporcionan la conectividad entre los equipos de usuario (UE) y la red troncal EPC. Un eNB se comunica con el resto de elementos del sistema mediante tres interfaces: E-UTRAN Uu, S1 y X2.

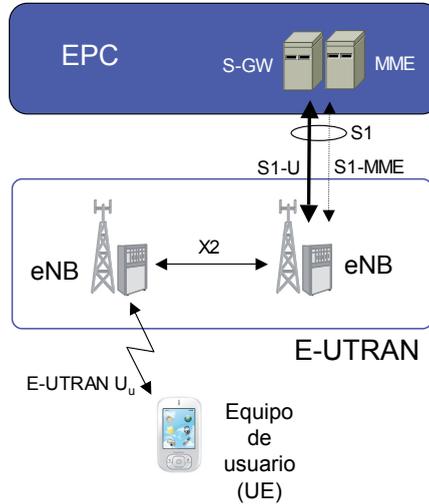


Figura 2.6 Red de acceso E-UTRAN

La interfaz E-UTRAN Uu, también denominada LTE Uu o simplemente interfaz radio LTE, permite la transferencia de información por el canal radio entre el eNB y los equipos de usuario. Todas las funciones y protocolos necesarios para realizar el envío de datos y controlar la operativa de la interfaz E-UTRAN Uu se implementan en el eNB.

El eNB se conecta a la red troncal EPC a través de la interfaz S1. Dicha interfaz está desdoblada en realidad en dos interfaces diferentes: S1-MME para sustentar el plano de control y S1-U como soporte del plano de usuario. Tal como se verá más detalladamente en el apartado 2.5.3, la separación entre plano de control y plano de usuario es una característica importante en la organización de las torres de protocolos asociadas a las interfaces de la red LTE. Así pues, el plano de usuario de una interfaz se refiere a la torre de protocolos empleada para el envío de tráfico de usuario a través de dicha interfaz (e.g., paquetes IP del usuario que se envían entre E-UTRAN y EPC a través de S1-U). Por otro lado, el plano de control se refiere a la torre de protocolos necesaria para sustentar las funciones y procedimientos necesarios para gestionar la operación de dicha interfaz o de la entidad correspondiente (e.g., configuración de la operativa del eNB desde la red EPC a través de S1-MME). Esta separación entre plano de control y plano de usuario en la interfaz S1 permite realizar la conexión del eNB con dos nodos diferentes de la red troncal. Así, mediante la interfaz S1-MME, el eNB se comunica con una entidad de red de la EPC encargada únicamente de sustentar las funciones relacionadas con el plano de control (dicha entidad de red de la red troncal EPC se denomina Mobility Management Entity, MME). Por otro lado, mediante la interfaz S1-U, el eNB se comunica con otra entidad de red encargada de procesar el plano de usuario (dicha entidad de red de la EPC se

denomina Serving Gateway, S-GW). Esta separación entre entidades de red dedicadas a sustentar el plano de control o bien el plano de usuario es una característica importante de la red LTE que permite dimensionar de forma independiente los recursos de transmisión necesarios para el soporte de la señalización del sistema y para el envío del tráfico de los usuarios.

Opcionalmente, los eNBs pueden conectarse entre si mediante la interfaz X2. A través de esta interfaz, los eNB se intercambian tanto mensajes de señalización destinados a permitir una gestión más eficiente del uso de los recursos radio (e.g., información para reducir interferencias entre eNBs) así como tráfico de los usuarios del sistema cuando estos se desplazan de un eNB a otro durante un proceso de *handover*.

En la Tabla 2.1 se resumen las entidades de red e interfaces de E-UTRAN y se indican las principales especificaciones del 3GPP relacionadas con cada una de ellas. La descripción detallada de las entidades de red e interfaces de E-UTRAN se aborda en el apartado 2.5.2.

Tabla 2.1 Entidades de red e interfaces de E-UTRAN

Entidades de red	Denominación	Descripción	Referencias 3GPP ¹⁾
	Evolved NodeB (eNB)	Estación base de la red de acceso E-UTRAN	TS 36.300 [4] TS 36.401 [5]
Interfaces	Denominación	Entidades de red asociadas	
	E-UTRAN Uu (también denominada LTE Uu o interfaz radio)	eNB UE	TS 36.300 [4] Documentos TS 36.2xx Y TS 36.3xx
	X2	eNB eNB	Documentos TS 36.42x TS 29.281 [26]
	S1-MME	eNB Red troncal EPC (MME)	Documentos TS 36.41x
	S1-U	eNB Red troncal EPC (S-GW)	TS 29.281 [26]

1) Las referencias proporcionadas son las más representativas aunque no constituyen una lista exhaustiva de todos los documentos 3GPP relacionados con cada aspecto.

2.5.2 Entidades de red e interfaces

En este apartado se describen más detalladamente las funciones asociadas a un eNB y las principales características de las tres interfaces de E-UTRAN.

2.5.2.1 Evolved NodeB (eNB)

Tal como se ha comentado en la descripción general de la arquitectura de E-UTRAN, el eNB integra todas las funciones de la red de acceso. Por ello, en el eNB terminan todos los protocolos específicos de la interfaz radio. Mediante dichos protocolos, el eNB realiza la transmisión de los paquetes IP hacia/desde los equipos de usuario junto con los mensajes de señalización necesarios para controlar la operación de la interfaz radio. El servicio de transferencia de paquetes IP entre un eNB y un equipo de usuario se denomina formalmente como servicio portador radio (*Radio Bearer*, RB). El eNB mantiene un *contexto* de cada uno de los equipos de usuario que tiene conectados. En dicho contexto se almacena la información necesaria para mantener los servicios de E-UTRAN activos (información sobre el estado del equipo de usuario, servicios portadores activos, información de seguridad, capacidades del terminal, etc.).

Sin duda, la funcionalidad clave de un eNB consiste en la gestión de los recursos radio. Así, el eNB alberga funciones de control de admisión de los servicios portadores radio, control de movilidad (p.ej, decisión de realizar un *handover*), asignación dinámica de los recursos radio tanto en el enlace ascendente como descendente (denominadas funciones de *scheduling*), control de interferencias entre estaciones base, control de la realización y del envío de medidas desde los equipos de usuario que puedan ser útiles en la gestión de recursos, etc. En el Capítulo 6 se proporciona una descripción detallada de las diferentes funciones de gestión de recursos radio necesarias para la operación eficiente de E-UTRAN.

Otra función importante introducida en la funcionalidad de un eNB es la selección dinámica de la entidad MME de la red troncal EPC cuando un terminal se registra en la red LTE. Esta función otorga un grado de flexibilidad muy importante en la operativa de la red. En E-UTRAN, a diferencia de arquitecturas más jerarquizadas como GERAN o las primeras versiones de UTRAN, un eNB puede estar conectado simultáneamente a múltiples MMEs de la red troncal. El conjunto de MMEs a los que tiene acceso un NB se denomina su *pool area*. Así, mediante la selección de qué entidad MME va a controlar el acceso de cada usuario, es posible balancear la carga de señalización entre diferentes MMEs así como aumentar la robustez del sistema frente a puntos de fallo críticos. Esta opción se soporta mediante lo que se denomina la interfaz S1 flexible (S1-flex) que se comentará más detalladamente en el apartado 2.5.2.3.

Al igual que la posibilidad de interactuar con múltiples MMEs, un eNB puede enviar/recibir paquetes IP de los usuarios a los que sirve a través de diferentes pasarelas S-GW de la red troncal EPC. Ello conlleva que el eNB albergue funciones de encaminamiento del tráfico de los usuarios hacia la pasarela de red S-GW correspondiente. La elección de S-GW en este caso compete a la entidad MME y no al eNB.

Un eNB puede gestionar una o varias celdas. Un caso típico es el uso de sectorización de forma que, el eNB ubicado en un emplazamiento soporta tantas celdas como sectores.

2.5.2.2 Interfaz radio

La interfaz radio soporta básicamente tres tipos de mecanismos de transferencia de la información en el canal radio: difusión de señalización de control, envío de paquetes IP y transferencia de señalización de control dedicada entre un equipo de usuario y el eNB. Los tres mecanismos citados se ilustran en la Figura 2.7 y se describen a continuación:

- Difusión (*broadcast*) de señalización de control en la zona de cobertura de la celda. La información enviada permite a los equipos de usuario detectar la presencia del eNB y conocer sus parámetros básicos de operación (e.g., potencia máxima que pueden utilizar los equipos de usuario en la celda) así como la identidad de los operadores de red a los que puede accederse a través del eNB. La información difundida corresponde tanto a información específica de la red de acceso (denominada información del *access stratum*, AS) como de la red troncal (denominada información del *non access stratum*, NAS). La difusión de señalización de control también sirve para forzar que un equipo de usuario que no tenga una conexión de control establecida con el eNB, inicie un acceso a la red (función de aviso o *paging*).
- Transferencia de paquetes IP de los usuarios a través del canal radio. Tal como se ha comentado anteriormente, los servicios de transferencia entre un eNB y un equipo de usuario se denominan servicios portadores radio (*Radio Bearers*, RB). Es importante destacar que los servicios portadores radio de E-UTRAN han sido diseñados específicamente para soportar tráfico IP y no permiten la transferencia de otros protocolos (e.g., paquetes X.25, tramas Ethernet, etc.). Por ello, de cara a la optimización del envío de tráfico IP a través de la interfaz radio, los servicios portadores albergan funciones como la compresión de cabeceras de los paquetes IP que permiten reducir el número de bytes enviados por la interfaz radio (las cabeceras de los paquetes IP pertenecientes a un mismo tipo de tráfico contienen un gran número de parámetros idénticos, p.ej, direcciones origen y destino, por lo que no resulta necesario enviar todos los bytes de la cabecera IP en cada uno de los paquetes).
- Transferencia de señalización de control dedicada entre el eNB y un equipo de usuario. El establecimiento de una conexión de control dedicada resulta imprescindible de cara a poder gestionar el uso de los servicios portadores radio así como para realizar cualquier gestión de señalización con la red troncal (e.g., registro del terminal en la red). La conexión de control se soporta mediante el protocolo *Radio Resource Control* (RRC). A través de dicho protocolo se gestionan, además del establecimiento, modificación y liberación de los servicios portadores radio entre el eNB y el equipo de usuario, otros mecanismos claves para la gestión eficiente de los recursos radio. Entre dichos mecanismos cabe citar el control y envío de medidas radio desde los terminales hacia el eNB y el mecanismo de *handover*, que permite que un equipo de usuario cambie de celda manteniendo activos tanto la conexión de control como los posibles servicios portadores radio que esté utilizando. Los terminales que mantienen una conexión de control con E-UTRAN se dice que se encuentran en modo conectado o activo, en con-

traposición al denominado modo *idle* en que el terminal no tiene una conexión RRC y básicamente se encuentra monitorizando la información de control difundida por la red. Una descripción más detallada de los estados de movilidad se realiza en el apartado 3.3.1 del Capítulo 3.

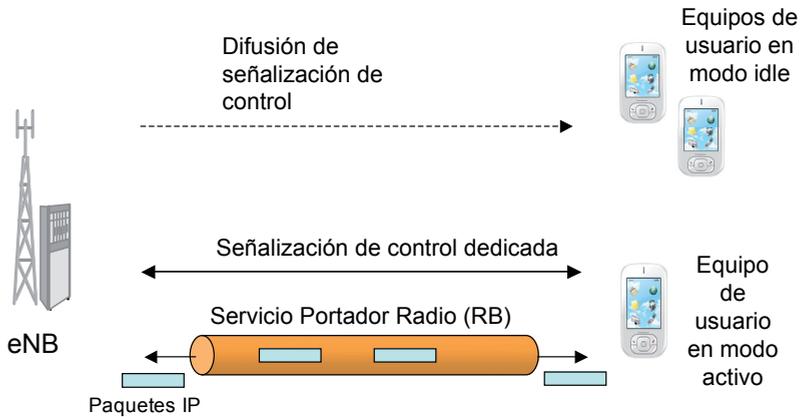


Figura 2.7 Ilustración de los mecanismos de transferencia de información en la interfaz radio

Respecto al envío de paquetes de usuario, cada servicio portador tiene asociado un perfil de QoS que debe satisfacerse mediante la correcta configuración de los protocolos radio así como la adecuada operación de los mecanismos de gestión de recursos radio (e.g., *scheduling*). En el apartado 3.2.3 del Capítulo 3 se proporcionan más detalles del modelo de QoS del sistema LTE.

La información enviada por la interfaz radio puede protegerse mediante funciones de cifrado que proporcionen confidencialidad e integridad. El servicio de confidencialidad permite que la información sea enviada de forma que ningún otro equipo que decodifique la señal transmitida por el canal radio sea capaz de conocer la información en claro. El servicio de integridad evita que la información transmitida pueda ser alterada de forma malintencionada en el camino entre eNB y equipo de usuario (e.g., equipo radio que se ubique en medio de la transmisión). Las funciones de cifrado se aplican tanto al tráfico de usuario (paquetes IP) como a los mensajes de señalización RRC utilizados en la conexión de control dedicada. Las claves secretas necesarias para realizar las funciones de seguridad se obtienen a través de los mecanismos de autenticación y gestión de claves abordados en el apartado 3.4 del Capítulo 3.

Una descripción más detallada de las funcionalidades de la interfaz radio en E-UTRAN se aborda en el Capítulo 5.

2.5.2.3 Interfaz eNB ↔ EPC (S1)

El plano de usuario de esta interfaz, denominado S1-U (*S1 User Plane*), proporciona un servicio de transferencia de datos de usuario entre eNB y S-GW sin garantías de entrega (se basa en UDP tal como se verá en el apartado 2.5.3.2 que detalla la torre de protocolos utilizada) y que no soporta ni mecanismos de control de errores ni de control de flujo. Este servicio de transferencia a través de la interfaz S1-U se denomina servicio portador S1 (*S1 bearer*).

El plano de control, denominado S1-MME o también S1-C, se utiliza para soportar un conjunto de funciones y procedimientos de control entre eNBs y la entidad MME de la red troncal. Concretamente, entre los procedimientos soportados en la interfaz S1 destacan:

- Procedimientos para establecimiento, modificación y liberación de recursos de los servicios portadores tanto en la interfaz radio (servicio portador radio o RB) como en la interfaz S1 (*S1 bearer*). La concatenación de un servicio portador radio y un servicio portador S1 constituye el servicio portador completo que ofrece la red de acceso E-UTRAN y que se denomina E-RAB (*E-UTRAN Radio Access Bearer*). Es importante tener en cuenta que en LTE, el establecimiento de estos servicios portadores que constituyen el plano de usuario para la transferencia del tráfico IP se controla desde la red troncal, en particular desde la entidad de red MME. Por tanto, en LTE no se permite que un eNB o el propio equipo de usuario puedan iniciar por su cuenta el establecimiento de un servicio portador radio. En la Figura 2.8 se ilustra dicho control del plano de usuario por parte de la entidad MME. El modelo de gestión de sesiones en la red EPC se analiza más detalladamente en el apartado 3.2 del Capítulo 3.
- Procedimientos de *handover* entre eNBs. Si la red E-UTRAN decide que un terminal debe cambiar de eNB en el transcurso de una conexión, y no existe una interfaz X2 entre los dos eNBs involucrados, la interfaz S1-MME se utiliza para articular el procedimiento de *handover*. De esta forma, a través de la interfaz S1-MME, la entidad MME puede establecer un nuevo contexto en el eNB destino asociado al terminal que va a realizar el cambio con toda la información relativa a la configuración de los servicios portadores que tiene establecidos el usuario así como las claves de seguridad. De esta forma, el re-establecimiento del servicio a través del nuevo eNB puede hacerse mucho más rápidamente ya que se evita el tener que ejecutar de nuevo los mecanismos para el establecimiento de los servicios portadores en la interfaz radio así como los mecanismos de seguridad. En el apartado 3.3.4 del Capítulo 3 se describe en detalle un procedimiento de *handover* basado en la utilización de la interfaz S1.
- Procedimiento de aviso (*Paging*). Tal como se verá más adelante en el apartado 2.6.2.1, una de las funciones básicas de la entidad MME es la gestión de la localización de los equipos de usuario en la red. La gestión de localización permite conocer con cierta resolución en qué eNB o conjunto de eNBs (denominados áreas de seguimiento, *Tracking areas*) puede ser localizado un usuario que se encuentre en modo *idle*, es decir, que no tenga establecida una conexión de control RRC con ningún eNB. Por ello, cuan-

do el MME quiere forzar que un usuario en modo *idle* pase a modo activo, a través de la interfaz S1-MME se ordena la ejecución del mecanismo de aviso en todos los posibles eNBs en los que espera encontrar al terminal.

- Procedimiento de envío de forma transparente entre MME y eNB de los mensajes de señalización de control que fluyen entre el MME y el equipo de usuario. Dichos mensajes corresponden a los protocolos denominados como protocolos NAS (*Non Access Stratrum*), cuya funcionalidad se verá más en detalle en el apartado 2.6.3.5.

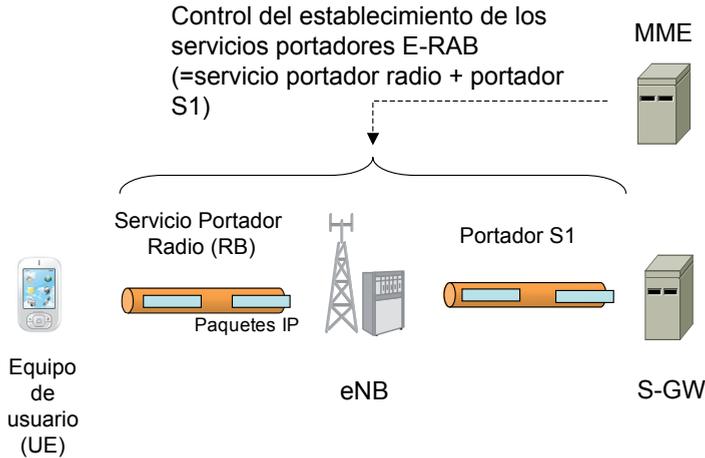


Figura 2.8 Control de los servicios portadores radio y S1 a través de la interfaz S1-MME

Tal como se ha comentado en la descripción de las funciones de un eNB, la interfaz S1 permite que un eNB pueda estar conectado simultáneamente a múltiples equipos de la red troncal (múltiples MMEs y/o múltiples S-GWs). Esta característica se denomina comúnmente como S1-flex y presenta varias ventajas:

- La red es más robusta a fallos de los nodos de la red troncal: el correcto funcionamiento de un eNB no está ligado a un único nodo de la red troncal. De esta forma, aunque se produzca un malfuncionamiento o parada de un nodo de la red troncal, los eNBs afectados pueden seguir ofreciendo servicio a través de otros nodos de la red troncal.
- Un eNB puede proporcionar acceso a nodos de redes troncales de diferentes operadores de red que comparten la red de acceso.
- Tal como se verá en el apartado 3.3 del Capítulo 3, cada terminal registrado en la red LTE tiene asociado un nodo MME que realiza el "seguimiento" del estado de dicho terminal. La frecuencia a la que debe cambiarse el nodo MME que controla el servicio de un determinado equipo de usuario puede verse reducida gracias a la posibilidad de seguir accediendo al mismo MME desde un conjunto mayor de eNBs. Esta reducción

del número de cambios de MME conlleva una reducción de la señalización en la red troncal necesaria para transferir la información de los usuarios de la base de datos HSS a la entidad MME que les da servicio.

- Es posible desplegar estrategias de balanceo de carga entre E-UTRAN y los nodos de la red troncal, atendiendo tanto a las capacidades de procesamiento de dichos nodos como a la capacidad de los enlaces de la red de transporte entre E-UTRAN y la EPC.

2.5.2.4 Interfaz eNB ↔ eNB (X2)

Al igual que el plano de usuario de S1, el plano de usuario de la interfaz X2 proporciona un servicio de transferencia de datos de usuario entre eNBs sin garantías de entrega y sin soporte de mecanismos de control de errores y de control de flujo. La transferencia de datos de usuario entre eNBs se realiza únicamente durante los procedimientos de *handover* en los que los paquetes de usuario almacenados en el eNB antiguo se transfieren al eNB nuevo. De esta forma, el cambio de eNB asociado a un procedimiento de *handover* puede resultar más transparente al usuario ya que se reduce la posible pérdida de paquetes durante el proceso. Nótese que, sobretodo en servicios de datos, el eNB antiguo podría tener acumulados en su buffer de transmisión paquetes IP del usuario en el momento del cambio. Dichos paquetes, cuando el usuario deja de estar accesible a través del eNB antiguo, podrían simplemente descartarse, con la consiguiente penalización en retardo y posible reducción en la tasa de transferencia del servicio asociado ya que la recuperación de dicha información recaería en la operación de las capas superiores (e.g., protocolo TCP en la capa de transporte). En cambio, si la propia red es capaz de transferir los paquetes IP del eNB antiguo al eNB nuevo, el impacto en el servicio puede reducirse notablemente.

Respecto al plano de control, entre las funciones y procedimientos soportados en la interfaz X2 destacan:

- Soporte del mecanismo de *handover* entre eNBs. En concreto, a través del plano de control se realiza la transferencia del contexto de un usuario del eNB antiguo al nuevo y se controla el mecanismo de transferencia de paquetes IP en el plano de usuario de X2. El contexto de usuario contiene información relativa a los servicios portadores radio que tiene establecidos el usuario, claves de seguridad así como los datos sobre las capacidades del terminal.
- Indicación del estado de carga del eNB. A través de dicha interfaz, eNBs que tengan celdas vecinas pueden transferirse información para llevar a cabo funciones de gestión de recursos radio como la coordinación de interferencias entre celdas que operen en el mismo canal.

2.5.3 Protocolos

Las torres de protocolos utilizadas en las tres interfaces de E-UTRAN (radio, S1 y X2) se

estructuran en torno a un plano de usuario y un plano de control. El plano de usuario abarca los protocolos utilizados para el envío del tráfico (paquetes IP) correspondiente a los servicios a los que acceden los terminales a través de la red. El plano de control se refiere a los protocolos necesarios para sustentar las funciones y procedimientos en las diferentes interfaces. A continuación se detallan los protocolos utilizados en las diferentes interfaces.

2.5.3.1 Protocolos en la interfaz radio

El envío de paquetes IP entre el eNB y un equipo de usuario a través de la interfaz radio se sustenta en una torre de protocolos formada por una capa de enlace (o capa de nivel 2) y una capa física. La torre de protocolos utilizada se muestra en la Figura 2.9. La capa de enlace se desglosa a su vez en tres subcapas: *Packet Data Convergence Protocol* (PDCP), *Radio Link Control* (RLC) y *Medium Access Control* (MAC). Cada capa/subcapa de la torre de protocolos se ocupa de un conjunto de funciones concreto y define el formato de los paquetes de datos (e.g., cabeceras y colas) que se intercambian entre entidades remotas. A continuación se describen las principales características de las diferentes capas/subcapas:

- *Packet Data Convergence Protocol* (PDCP). Constituye la capa superior de la torre de protocolos encargada de proporcionar el punto de acceso al servicio portador radio (*Radio Bearer*, RB). Es decir, los paquetes IP del tráfico de usuario se entregan y se reciben a través del servicio de transferencia proporcionado por la capa PDCP. Las funciones principales de esta capa son la compresión de cabeceras de los paquetes IP y el cifrado de la información para garantizar su confidencialidad e integridad. La cabecera añadida por la capa PDCP básicamente contiene un número de secuencia que identifica al paquete IP enviado y permite realizar una entrega ordenada de los paquetes IP en el extremo receptor así como detectar posibles duplicados de los paquetes IP (ocasionados por ejemplo en un proceso de *handover*). Cada servicio portador radio tiene una entidad PDCP asociada.
- *Radio Link Control* (RLC). La capa RLC permite enviar de forma fiable los paquetes PDCP entre el eNB y equipo de usuario. Para ello, la capa RLC soporta funciones de corrección de errores mediante mecanismos *Automatic Repeat ReQuest* (ARQ), concatenación, segmentación y re-ensamblado, entrega ordenada de paquetes PDCP a capas superiores (excepto durante el mecanismo de *handover*), detección de duplicados y detección/recuperación de errores en el protocolo. Cada servicio portador radio tiene una entidad RLC asociada.
- *Medium Access Control* (MAC). Es la capa encargada de controlar el acceso al canal radio. Para ello, la capa MAC soporta funciones de *scheduling* dinámico entre equipos de usuario atendiendo a prioridades, multiplexa los paquetes RLC de diferentes servicios portadores radio en los canales de transporte ofrecidos por la capa física (un canal de transporte puede ser compartido por varios servicios portadores de uno o varios equipos de usuario) y realiza un control de errores mediante *Hybrid ARQ* (HARQ). Los servicios de transferencia que la capa MAC ofrece a la capa RLC se denominan canales lógicos. Existe una única entidad MAC por celda.

- **Capa física.** Es la capa encargada de realizar la transmisión propiamente dicha a través del canal radio. Alberga funciones de codificación de canal, modulación, procesamiento asociado a las técnicas de múltiples antenas de transmisión/recepción, y mapeo de la señal a los recursos físicos frecuencia-tiempo apropiados. En el enlace ascendente, la capa física se basa en un esquema *single-carrier FDMA*. En el enlace descendente, el esquema de transmisión es *OFDMA*. En el Capítulo 4 se analizan los fundamentos básicos de los mecanismos de transmisión utilizados en la capa física de LTE. Los servicios de transferencia que la capa física ofrece a la capa MAC se denominan canales de transporte. Existe una única entidad de capa física por celda.

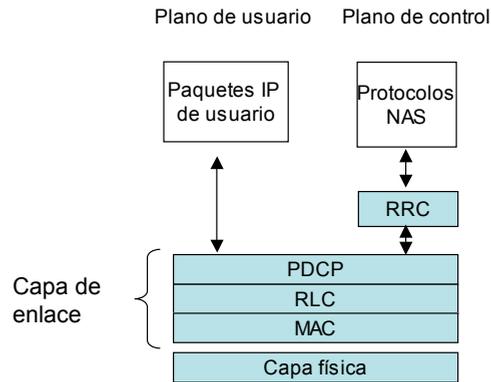


Figura 2.9 Protocolos de la interfaz radio de E-UTRAN

Respecto al plano de control entre el equipo de usuario y la red, éste se soporta sobre la misma capa de enlace (protocolos PDCP, RLC, MAC) y la misma capa física utilizadas en el plano de usuario. Los protocolos de nivel de red específicos de este plano son:

- **Radio Resource Control (RRC).** Esta capa permite establecer una conexión de control entre el eNB y un equipo de usuario a través de la cual se llevan a cabo un número importante de funciones relacionadas con la gestión de la operativa de la interfaz radio. Entre dichas funciones de la capa RRC destacan los mecanismos de gestión de los servicios portadores radio (e.g., señalización para el establecimiento/liberación/modificación de los portadores radio), el soporte de funciones de movilidad (e.g., señalización de *handover*), la difusión (*broadcast*) de parámetros de sistema y funciones de aviso de los terminales que no disponen de una conexión RRC establecida (e.g., envío de avisos a través del canal de *paging*). El servicio de transferencia que ofrece la capa PDCP para el envío de los mensajes de señalización del protocolo RRC se denomina servicio portador de señalización (*Signalling Radio Bearer, SRB*).
- **Señalización de los protocolos NAS.** Los protocolos NAS se extienden entre la entidad de red MME en la red troncal y el equipo de usuario. Los mensajes de estos protocolos se transportan de forma transparente en la interfaz radio encapsulados dentro de la parte de datos de los mensajes RRC. Las principales funciones de los protocolos NAS

son: autenticación, autorización, gestión de movilidad de los terminales que no tienen una conexión RRC establecida y gestión de los servicios portadores de la red EPS.

En el Capítulo 5 se analizan más en detalle los protocolos utilizados en la interfaz radio de LTE (RRC, PDCP/RLC/MAC), así como las características de la capa física de E-UTRAN. Por otro lado, los protocolos NAS se discuten más detalladamente en el apartado 2.6.3.5 de este mismo capítulo.

2.5.3.2 Protocolos en las interfaces S1 y X2

La estructura de protocolos utilizada en E-UTRAN para soportar las interfaces S1 y X2 establece una separación entre la capa de red radio (*Radio Network Layer, RNL*) y la capa de red de transporte (*Transport Network Layer, TNL*), tal como ya introdujo la red UMTS. Esta descomposición tiene como objetivo aislar las funciones que son específicas del sistema de comunicaciones móviles (UMTS o LTE), de aquellas otras que dependen de la tecnología de transporte utilizada (e.g., IP, ATM). De esta forma, los protocolos específicos de la red de acceso radio constituyen la capa RNL mientras que la capa TNL alberga los protocolos utilizados para el transporte de la información de la capa RNL entre las entidades de la red. En la Figura 2.10 se ilustra la arquitectura de protocolos de las interfaces S1 y X2. La separación entre las capas TNL y RNL en E-UTRAN se establece en el documento [5]. Los documentos de base del 3GPP que especifican la estructura de las interfaces S1 y X2 son, respectivamente, 3GPP TS 36.410 [6] y TS 36.420 [7].

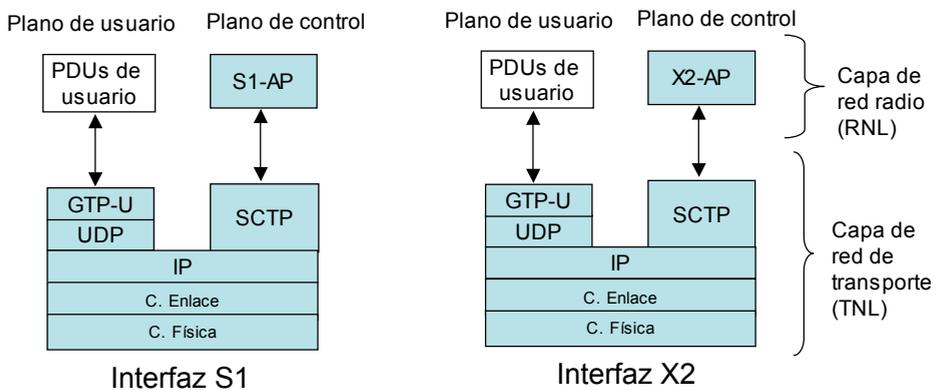


Figura 2.10 Protocolos en las interfaces S1 (izquierda) y X2 (derecha)

Tanto el plano de usuario de la interfaz S1 (S1-U) como el de la interfaz X2 utilizan el protocolo de encapsulado GTP-U (*GPRS Tunneling Protocol – User Plane*) para el envío de paquetes IP de usuario. El protocolo GTP-U es un protocolo heredado de GPRS que en las redes GSM y UMTS se utiliza dentro del dominio de paquetes de la red troncal (e.g., en la interfaz entre SGSN y GGSN) así como en el plano de usuario de la interfaz lu-PS de la red de acceso

UTRAN. En las interfaces S1-U y X2, el protocolo GTP-U se transporta sobre UDP/IP y fundamentalmente se utiliza para multiplexar los paquetes IP de múltiples usuarios (los paquetes IP de un determinado servicio portador se encapsulan con una determinada etiqueta –identificador de tunel—). En el apartado 2.6.3.1 se proporcionan detalles adicionales del protocolo GTP-U, dada su gran relevancia en el contexto de las redes 3GPP. Finalmente, es importante destacar que los planos de usuario de ambas interfaces no contemplan mecanismos de entrega garantizada para la transferencia de los paquetes de usuario, ni tampoco mecanismos de control de errores o control de flujo.

Respecto al plano de control de la interfaz S1 (S1-MME o S1-C), la capa de red radio consiste en el protocolo S1-AP (*S1 - Application Part*). Este protocolo es el que sustenta los procedimientos soportados en la interfaz S1 comentados en el apartado 2.5.2.3 (establecimiento de servicios portadores en el eNB, control del *handover*, *paging*, etc.). La especificación del protocolo se realiza en el documento 3GPP TS 36.413 [8]. La transferencia de los mensajes de señalización del protocolo S1-AP entre eNBs y MMEs se realiza mediante el servicio de transferencia fiable que ofrece el protocolo de transporte *Stream Control Transmission Protocol* (SCTP).

SCTP es un protocolo de transporte (al igual que otros protocolos como TCP y UDP) de propósito general estandarizado por IETF en la RFC4960 [9] que fue concebido originariamente para el envío de señalización de redes telefónicas sobre redes IP. SCTP hereda muchas de las funciones contempladas en TCP a la vez que introduce importantes mejoras encaminadas a proporcionar mayor robustez y versatilidad en la transferencia de diferentes tipos de información. En particular, al igual que TCP, SCTP dispone de mecanismos de control de flujo y de congestión en la conexión, denominada asociación en SCTP. Por otro lado, SCTP incorpora soporte para *multihoming* (las asociaciones soportan la transferencia a través de múltiples caminos entre los nodos participantes, es decir, los nodos participantes pueden disponer de múltiples direcciones IP), *multi-streaming* (múltiples flujos pueden enviarse en paralelo en el seno de una misma asociación) y el envío de la información se estructura en base a mensajes (a diferencia del protocolo TCP que trata la información como una secuencia de bytes). Estas nuevas capacidades son las que hicieron que en 3GPP se optara por la utilización de este protocolo, en lugar de TCP, para implementar el plano de control de las interfaces S1 y X2 de E-UTRAN.

Respecto al plano de control de la interfaz X2, el protocolo utilizado para sustentar los procedimientos descritos en el apartado 2.5.2.4 se denomina X2-AP (X2 Application Part) y se especifica en 3GPP TS 36.423 [10].

2.5.3.3 Plano de usuario entre UE y EPC

Atendiendo a la descripción realizada en los anteriores apartados, en la Figura 2.11 se ilustra el plano de usuario completo de E-UTRAN para el envío de paquetes IP entre el equipo de usuario (UE) y la red troncal (S-GW). Los paquetes IP contienen la información correspondiente al servicio que el usuario está utilizando (voz, video, datos) así como la señalización a nivel de aplicación (protocolos SIP, RTCP, etc.). El eNB realiza funciones de “*relay*” entre la torre de

protocolos PDCP/RLC/MAC/PHY de la interfaz radio y la torre de protocolos de la interfaz S1-U. Es importante destacar que el eNB no realiza ninguna decisión de encaminamiento a partir de la información contenida en las cabeceras IP de los paquetes de usuario sino que simplemente se ocupa de su transferencia entre las dos interfaces atendiendo a los servicios portadores establecidos. La estructura y encaminamiento de la información en la red LTE basada en servicios portadores se cubre en el apartado 3.2 del Capítulo 3.

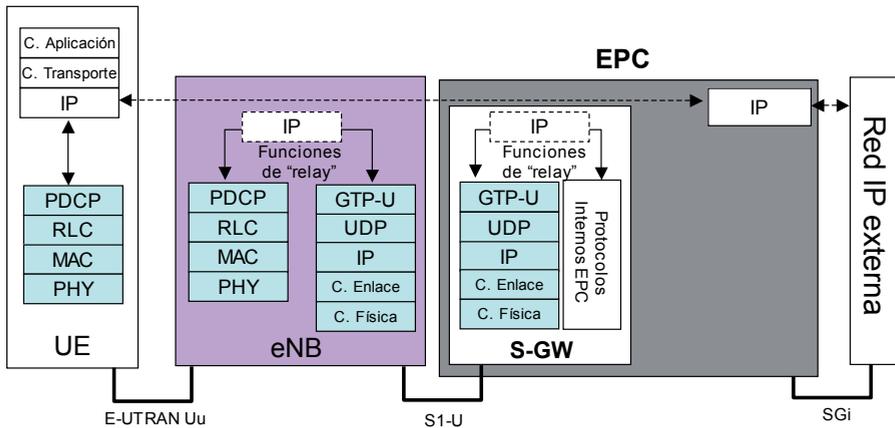


Figura 2.11 Protocolos del plano de usuario en E-UTRAN

2.5.3.4 Plano de control entre UE y EPC

En la Figura 2.12 se ilustra la torre de protocolos del plano de control para el envío de señalización NAS entre el equipo de usuario y la red troncal. Los protocolos NAS (descritos en el apartado 2.6.3.5) se transportan encapsulados (de forma transparente) dentro de mensajes RRC en la interfaz radio y en mensajes S1-AP en la interfaz S1-MME. El eNB realiza las funciones de "relay" necesarias entre ambas torres de protocolos.

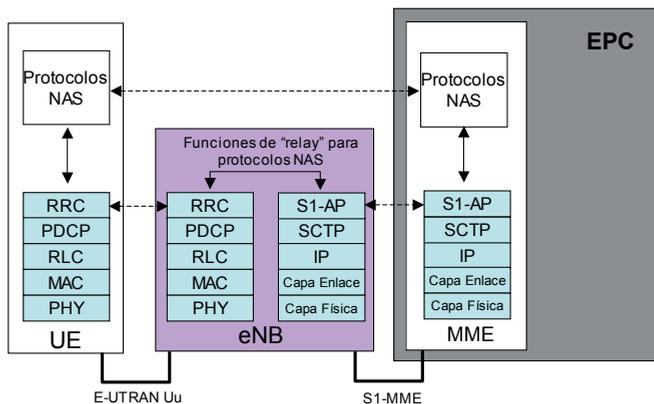


Figura 2.12 Protocolos del plano de control en E-UTRAN

2.5.4 Comparativa E-UTRAN y UTRAN

La arquitectura E-UTRAN presenta importantes diferencias con respecto a las redes de acceso UTRAN y GERAN. En la Figura 2.13 se muestra a nivel ilustrativo un despliegue simple de una red E-UTRAN y de una red UTRAN, como ejemplo típico tanto de redes de acceso 3G como 2G, en aras a comparar ambas arquitecturas.

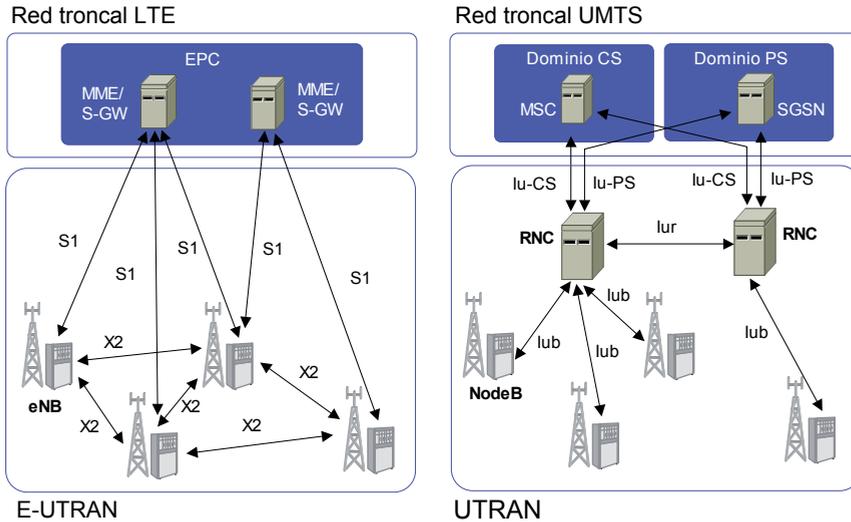


Figura 2.13 Comparativa de la arquitectura de red de E-UTRAN y UTRAN

Las redes de acceso anteriores a E-UTRAN se basan en una arquitectura jerárquica donde las funciones de la red de acceso se distribuyen en dos tipos de nodos: estaciones base (denominados Nodos B en UTRAN) y equipos controladores de estas estaciones base (denominados RNC en UTRAN). La arquitectura de UTRAN se especifica en el documento 3GPP TS 25.401 [11]. En esta arquitectura jerarquizada, los equipos controladores albergan el plano de control de la interfaz radio (señalización de control del enlace radio) así como múltiples funciones del plano de usuario (algunas funciones de la capa de acceso al medio, control de enlace, compresión de cabeceras, etc.). Por otro lado, las estaciones base se ocupan principalmente de las funciones de transmisión radio (procesado de capa física) y su operación se gestiona de forma remota desde los equipos controladores. La interconexión entre estaciones base y controladores se realiza mediante una interfaz denominada lub de forma que la topología de red resultante a nivel lógico es una topología en forma de estrella. Los equipos controladores también pueden conectarse entre si mediante interfaces específicas como la interfaz lur que, en el caso de UTRAN, permite la explotación del mecanismo de *macrodiversidad* entre dos Nodos B que se encuentren conectados a RNCs diferentes. La interconexión de la red de acceso a la troncal se realiza a través de los equipos controladores mediante las interfaces lu-PS, entre RNCs y los nodos SGSNs del dominio de paquetes, y lu-CS, entre RNCs y las centrales de conmutación MSC del dominio de circuitos.

Comparando la arquitectura de UTRAN con E-UTRAN, puede observarse en la Figura 2.13 que E-UTRAN sigue una arquitectura “plana”, sin ningún nivel de jerarquización. Tal como se ha indicado en la lista de funciones asociadas a un eNB, los protocolos radio se ejecutan íntegramente en los eNBs (no es necesario ningún equipo adicional como el RNC de UTRAN). Es importante destacar que la integración de los protocolos radio de capa física y de enlace en la estación base es una característica adoptada también en otras interfaces radio como IEEE 802.11 para redes de área local y IEEE 802.16 utilizada en WiMAX. La interconexión de E-UTRAN con la red troncal se realiza en cada uno de las estaciones base (eNBs) mediante la interfaz S1. Tal como se ha comentado anteriormente, la interfaz S1 soporta configuraciones donde un eNB puede estar conectado simultáneamente con múltiples elementos de la EPC (varios MME y/o varios S-GW). Esto hace que el dimensionamiento de la red de acceso (eNBs) y de los equipos de la red troncal (MME y pasarelas S-GW) pueda hacerse de forma más flexible, permitiendo, por ejemplo, que el tráfico cursado a través de los eNBs se derive hacia el nodo de la red troncal más adecuado atendiendo a criterios de balanceo de cargas. Por el contrario, en una estructura jerárquica en árbol como la utilizada en UTRAN, la capacidad sobrante en nodos ubicados en ramas diferentes no puede ser aprovechada. Asimismo, aunque de forma opcional, las estaciones base de E-UTRAN pueden conectarse directamente entre si formando una topología semi-mallada (un eNB puede conectarse a un subconjunto de eNBs mediante la interfaz X2) que permite tanto la transferencia de información de control como de tráfico de usuario entre ellas. Esta opción no está contemplada en UTRAN (los Nodos B no se interconectan entre ellos).

Uno de los principales motivos que condujeron a la utilización de arquitecturas jerárquicas en los sistemas 2G y 3G fue básicamente económico: la concentración de los recursos de procesado en unos pocos equipos capaces de servir a un elevado conjunto de usuarios a través de estaciones base poco complejas, y por tanto, con un coste “relativamente” menor, constituía una opción más competitiva en términos de coste frente a la alternativa de utilizar estaciones base mucho más complejas capaces de albergar la mayoría de las funciones propias del sistema de comunicaciones [12] [13]. Este argumento ha ido perdiendo peso de forma progresiva en los últimos años conforme a los avances realizados en las tecnologías de computación que permiten disponer de plataformas de procesado muy potentes a costes reducidos. Adicionalmente, otros argumentos que han propiciado la transición a arquitecturas planas como E-UTRAN, en lugar de arquitecturas jerarquizadas, han sido la explotación de mecanismos de diversidad temporal como H-ARQ y *schedulers* rápidos que requieren ser ejecutados en la propia estación base para conseguir tiempos de ida y vuelta muy reducidos. La explotación de estos mecanismos de diversidad conlleva a la vez que no sea necesario soportar mecanismos de diversidad adicionales como la macrodiversidad (transmisión y recepción simultánea de un equipo de usuario en varias celdas) para mejorar las prestaciones del enlace radio. Nótese que el soporte de macrodiversidad es uno de los pilares fundamentales de UTRAN que requiere la realización de funciones específicas de combinación y bifurcación del tráfico dentro de la red de acceso, en particular, en los equipos controladores RNC. También, en términos de escalabilidad y robustez, existen argumentos a favor de las arquitecturas

planas frente a las centralizadas. La existencia de un elemento crítico y de alto coste como el controlador en las arquitecturas centralizadas condiciona la escalabilidad de la red de acceso (e.g., si se requiere aumentar la capacidad o cobertura de la red mediante la instalación de una estación base adicional, y el controlador ya se encuentra al límite de su capacidad, sería necesario introducir un nuevo controlador en la red simplemente para poder incorporar una nueva estación base). Además, un elemento controlador constituye un punto de fallo crítico que puede afectar al funcionamiento de muchas estaciones base y, por tanto, afectar a un elevado número de usuarios (un controlador RNC puede gestionar varios centenares de estaciones base).

2.6 Red troncal de paquetes evolucionada: EPC

2.6.1 Arquitectura de EPC

El diseño de la red troncal EPC ha sido concebido principalmente para proporcionar un servicio de conectividad IP (evolución del servicio GPRS) mediante una arquitectura de red optimizada que permite explotar las nuevas capacidades que ofrece la red de acceso E-UTRAN. Asimismo, otro factor clave considerado en el diseño de la arquitectura de la red troncal ha sido la posibilidad de acceder a sus servicios a través de otras redes de acceso tanto 3GPP (UTRAN y GERAN) como fuera del ámbito del 3GPP (cdma2000, WiMAX, 802.11). La descripción completa de la red troncal EPC se recoge en los documentos 3GPP TS 23.401 [14] y 3GPP TS 23.402 [15]. En particular, en la especificación TS 23.401 se cubre la arquitectura de la red troncal EPC cuando la red de acceso es E-UTRAN, así como la utilización de redes de acceso 3GPP alternativas o complementarias como UTRAN y GERAN. Por otro lado, la especificación TS 23.402 extiende la arquitectura de la red troncal EPC para soportar el acceso a través de otras redes no 3GPP.

De cara a introducir los diferentes componentes de la arquitectura completa de EPC de forma progresiva, la arquitectura mostrada en la Figura 2.14 comprende únicamente las entidades de red que forman el núcleo de la red troncal EPC para la provisión de servicios de conectividad IP a través de una red de acceso E-UTRAN, junto con las entidades de red e interfaces que soportan las funciones relacionadas con el control de del servicio de conectividad (e.g., control de QoS) y de los mecanismos de tarificación. El soporte de las redes de acceso UTRAN y GERAN, así como otras redes no 3GPP, se aborda en el Capítulo 7 a partir de las bases establecidas en este apartado.

Es importante matizar en este punto que las entidades de red en base a las cuales se realiza la descripción de la arquitectura de la red troncal son entidades funcionales: una entidad de red en 3GPP se concibe como una entidad "lógica" que cubre una funcionalidad perfectamente delimitada. Por tanto, una implementación concreta de la red troncal EPC admite que diferentes entidades funcionales puedan residir en el mismo equipo físico, tal como se analizará posteriormente en el apartado 2.6.4.

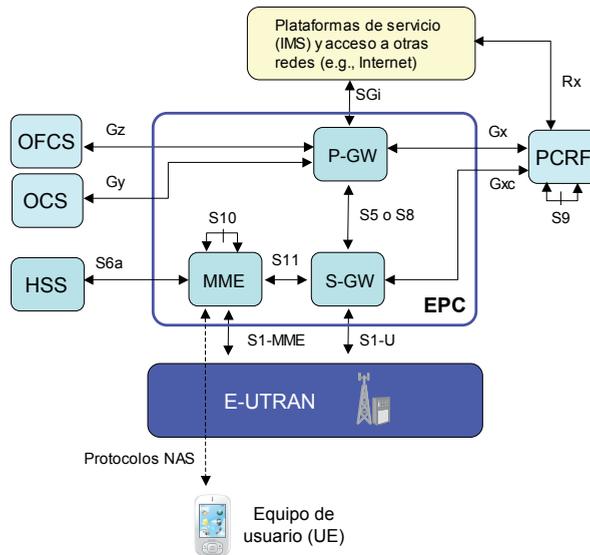


Figura 2.14 Arquitectura básica de la red troncal EPC

Tal como se ilustra en la Figura 2.14, el núcleo del sistema EPC está formado por tres entidades de red: MME (*Mobility Management Entity*), Serving Gateway (S-GW) y *Packet Data Network Gateway* (P-GW). Estas tres entidades, junto con la base de datos principal del sistema 3GPP denominada HSS (*Home Subscriber Server*), constituyen los elementos básicos para la provisión del servicio de conectividad IP entre los equipos de usuario conectados a través de E-UTRAN y redes externas a las que se conecta la red troncal EPC. Las funciones asociadas con el plano de usuario se concentran en las dos pasarelas (S-GW y P-GW) mientras que la entidad MME se encarga de las funciones y señalización del plano de control.

La interconexión de la red de acceso E-UTRAN a la EPC se realiza a través de la interfaz S1 (descrita más detalladamente en el apartado 2.5). En particular, la interfaz S1-MME que sustenta el plano de control termina en la entidad MME mientras que la interfaz S1-U del plano de usuario termina en el S-GW.

La entidad MME termina el plano de control de los equipos de usuario conectados a la red LTE mediante los protocolos NAS y controla las funciones de transferencia del plano de usuario de red LTE a través de la interfaz S11⁴ con la pasarela S-GW. Asimismo, la entidad MME se conecta a la entidad HSS a través de la interfaz S6a para acceder a la información asociada a los usuarios de la red que estén autorizados a establecer conexiones a través de E-UTRAN. Tal como puede observarse en la Figura 2.14, las entidades MME también pueden comunicarse entre ellas mediante la interfaz S10.

⁴ Las interfaces específicas de la red troncal EPC se han codificado con un prefijo formado por la letra “S” y unos dígitos y letras adicionales, de forma similar al criterio adoptado en su día para GPRS donde las interfaces empiezan por el prefijo “G”.

Por otro lado, la interconexión de la EPC con redes externas o plataformas de servicio (e.g., plataformas IMS) se realiza a través de la pasarela P-GW mediante la interfaz SGi. La pasarela P-GW soporta funciones, entre otras, de asignación de direcciones IP a los equipos de usuario y mecanismos de control de los parámetros de calidad de servicio de las sesiones de datos establecidas a través de la red LTE. Internamente, la pasarela P-GW se conecta a la pasarela S-GW mediante la interfaz S5, cuando ambas pasarelas pertenecen al mismo operador, y mediante S8, cuando éstas se encuentran en redes de operadores diferentes y se proporciona un servicio de *roaming* o itinerancia (las configuraciones de red en escenarios de itinerancia se verán con más detalle en el apartado 2.6.5.).

También ilustrada en la Figura 2.14, la entidad de red PCRF (*Policy and Charging Rules Function*) constituye un elemento clave de todos los sistemas 3GPP, y en particular, del sistema LTE. La entidad PCRF forma parte del marco funcional denominado PCC (*Policy and Charging Control*) [27] que se utiliza para controlar los servicios portadores que ofrece la red LTE (e.g., activación y determinación de los parámetros de QoS asociados a cada servicio portador) así como realizar el control de los mecanismos de tarificación (e.g., tarificación on-line, offline, medición del volumen de datos transferido, tiempo transcurrido, etc.). Así pues, mediante la interfaz Gx, el PCRF gestiona los servicios portadores EPS de la red LTE mediante el envío de unas reglas de uso (i.e., reglas PCC) que sirven para configurar la operación de unas funciones específicas del plano de usuario de la pasarela P-GW (e.g., funciones que limitan la tasa de transferencia en bits/s de los servicios portadores). La entidad PCRF es accesible desde las plataformas de servicios externas como IMS mediante la interfaz Rx. Dicha interfaz ofrece la funcionalidad de control necesaria para que los servidores de aplicación externos puedan proporcionar información asociada a los servicios finales a los que accede el usuario junto con las características y requerimientos de QoS. A modo de ejemplo, si un usuario establece un servicio de videoconferencia a través de IMS, el elemento que controla la provisión del servicio en IMS puede indicar a través de la interfaz Rx cuáles son los parámetros de QoS que debe proporcionar el servicio portador de la red LTE para transferir de forma adecuada la información de la videoconferencia. Con esta información, la entidad PCRF envía a la red LTE las reglas PCC pertinentes para la configuración de los servicios portadores. Una descripción más detallada del marco de PCC adoptado en los sistemas 3GPP, junto con las entidades de red e interfaces asociadas, se aborda en el apartado 3.2.4 del Capítulo 3.

Finalmente, las entidades OFCS (*Offline Charging System*) y OCS (*Online Charging System*) mostradas en la Figura 2.14 constituyen el núcleo del sistema de tarificación de la red [28]. Ambas entidades interactúan directamente con la pasarela P-GW mediante la interfaz Gz, en el caso de OFCS, y Gy, en el caso de OCS. El marco de tarificación soportado es un marco flexible que permite desplegar modelos de tarificación en base a diferentes parámetros tales como tiempo de uso, volumen de datos, eventos, etc. Algunos detalles adicionales del sistema de tarificación relacionados con el marco PCC comentado anteriormente se proporcionan en el apartado 3.2.4 del Capítulo 3. Una descripción más detallada del sistema de tarificación queda fuera del ámbito de este libro.

En la Tabla 2.2 se resumen las entidades de red e interfaces propias de EPC junto con las otras entidades comunes de los sistemas 3GPP que han sido mencionadas en este apartado. En la tabla también se indican las especificaciones del 3GPP más relevantes relacionadas con cada una de ellas. En el siguiente apartado se proporcionan más detalles de las entidades de red e interfaces de la red troncal EPC.

Tabla 2.2 Entidades de red e interfaces de EPC para el acceso desde E-UTRAN

	Denominación	Descripción	Referencias 3GPP ¹⁾
Entidades de red EPC	MME	Nodo que canaliza el plano de control de la red LTE	TS 23.401 [5]
	S-GW	Punto de anclaje del plano de usuario en la red troncal	TS 23.401 [14]
	P-GW	Pasarela para la interconexión con redes externas	TS 23.401 [14]
Entidades comunes a las redes 3GPP	Denominación	Descripción	
	HSS	Base de datos global del sistema (contiene, entre otros, los datos de suscripción de los usuarios).	TS 23.002 [1] TS 23.008 [16]
	PCRF	Elemento central del sistema de control de uso de la red (policy control) y control de tarificación	TS 23.203 [27]
	OCS	Nodo de control para la provisión de servicios que requieren tarificación on-line	TS 23.203 [27] TS 32.240 [28]
	OFCS	Nodo de recogida de la información de tarificación para su posterior transferencia al sistema de facturación.	TS 23.203 [27] TS 32.240 [28]
Interfaces	Denominación	Entidades de red asociadas	
	S1-MME	MME E-UTRAN (eNB)	Documentos TS 36.41x
	S1-U	S-GW E-UTRAN (eNB)	TS 29.281 [26]
	SGi	P-GW Redes externas	TS 29.061 [17]
	S6a	MME HSS	TS 29.272 [25]
	S5/S8	P-GW S-GW	TS 29.274 [21] (opción GTP) TS 23.275 [24] (opción PMIPv6)
	S11	MME S-GW	TS 29.274 [21]
	S10	MME MME	TS 29.274 [21]
	Señalización NAS	UE MME	TS 24.301 [37] [37]
	Rx	PCRF Plataformas servicios	TS 29.214 [29]
	S9	PCRF PCRF	TS 29.215 [30]
	Gx/Gxc	P-GW/S-GW PCRF	TS 29.212 [31]
	Gz/Gy	P-GW OFCS/OCS	Documentos TS 32.2xx

¹⁾ Las referencias proporcionadas son las más representativas pero no constituyen una lista exhaustiva de todos los documentos 3GPP relacionados con cada aspecto.

2.6.2 Entidades de red e interfaces

En este apartado se describen más detalladamente las funciones soportadas por las tres entidades de red que forman el núcleo de la red troncal EPC (MME, S-GW y P-GW) y la entidad HSS común al resto de sistemas 3GPP, junto con las interfaces SGi, S5/S8, S11, S10 y S6a. Las interfaces S1-MME y S1-U se han abordado ya en el apartado 2.5 en el contexto de la red de acceso E-UTRAN. El resto de interfaces relacionadas con el sistema PCC, Rx/S9/Gx/Gxc/Gz/Gy, se cubren en el apartado 3.2.4 del Capítulo 3.

2.6.2.1 MME (Mobility Management Entity).

La entidad MME constituye el elemento principal del plano de control de la red LTE para gestionar el acceso de los terminales a través de E-UTRAN. Todo terminal que se encuentre registrado en la red LTE y sea accesible a través de E-UTRAN, tiene una entidad MME asignada. La elección de la entidad MME se realiza en el proceso de registro y depende de aspectos tales como la ubicación geográfica del terminal en la red (cada MME sirve a un conjunto determinado de eNBs) así como a criterios de balanceo de cargas (gracias al soporte de la interfaz S1-flex comentada en el apartado 2.5.2.3). Dicha entidad mantiene un contexto de datos del usuario (e.g., identificadores del usuario, conexiones y servicios portadores EPS activos, claves de seguridad, datos de localización del usuario en la red, etc.) y articula todas las gestiones que se realicen en relación a dicho usuario (e.g., establecimiento de servicios portadores EPS, etc.). La entidad MME asignada a un usuario puede ir cambiando atendiendo a la movilidad de dicho usuario dentro de la zona de servicio de la red. Las principales funciones de la entidad MME son las siguientes:

- Autenticación y autorización del acceso de los usuarios a través de E-UTRAN. A partir de los datos de usuario obtenidos desde el HSS, la entidad MME se encarga de llevar a cabo el control de acceso a la red mediante la identificación, autenticación y autorización de los usuarios que se conectan a través de E-UTRAN. Los detalles de las funciones y procedimientos del control de acceso se abordan en el apartado 3.4.4 del Capítulo 3.
- Gestión de los servicios portadores EPS. La entidad MME es la encargada de articular la señalización necesaria para establecer, mantener, modificar y liberar los servicios portadores EPS sobre los cuales se sustenta el envío de paquetes IP entre los equipos de usuario y la red externa.
- Gestión de movilidad de los usuarios en modo *idle* (i.e. terminales que no tienen ninguna conexión de control establecida con E-UTRAN). La entidad MME es la encargada de hacer un seguimiento de la localización de los usuarios dentro del área de servicio de la red. Para ello, se definen unas áreas de seguimiento (*Tracking areas*) y unos procedimientos asociados (denominados *Tracking Area Update*) que permiten disponer de información de localización de todos los usuarios que se encuentren registrados en la red LTE. Esta gestión, conocida específicamente como gestión de localización, se describe en el apartado 3.3.2 del Capítulo 3.

- Señalización para el soporte de movilidad entre EPS y redes 3GPP. Tal como se verá en el Capítulo 7, la entidad MME de la EPC y la entidad SGSN pueden intercambiarse información relativa a los equipos de usuario conectados bien a través de E-UTRAN o de UTRAN/GERAN para poder gestionar, por ejemplo, mecanismos de gestión de movilidad conjunta (la red troncal GPRS así como la red troncal EPC pueden intercambiar información relativa a las áreas de seguimiento). Dicha señalización se realiza a través de la interfaz S3 entre MME y SGSNs. También, a través de esta interfaz, se gestionan los procedimientos de reubicación del plano de usuario en las entidades de la red troncal (e.g., el plano de usuario de un terminal conectado inicialmente a UTRAN y que fluye a través de un determinado SGSN, se reubica hacia una pasarela S-GW cuando el terminal cambia de UTRAN a E-UTRAN).
- Terminación de los protocolos de señalización NAS (*Non Access Stratum*). Los protocolos NAS fluyen entre el equipo de usuario y la entidad MME que tenga asignada. A través de ellos se soportan los procedimientos relacionados con las funciones de control de acceso a la red LTE, la gestión de las conexiones a redes externas y el establecimiento de servicios portadores EPS, y la gestión de movilidad de los terminales que se encuentran en modo *idle*.

2.6.2.2 Serving Gateway (S-GW).

Esta entidad actúa de pasarela del plano de usuario entre E-UTRAN y la red troncal EPC. Al igual que sucede con la entidad MME, un usuario registrado en la red LTE dispone de una entidad S-GW asignada en la EPC a través de la cual transcurre su plano de usuario. La asignación de la pasarela S-GW responde también a criterios geográficos así como de balanceo de cargas. Entre las principales funciones del S-GW podemos destacar:

- La entidad S-GW proporciona un punto de anclaje en la red troncal EPC con respecto a la movilidad del terminal entre eNBs. De esta forma, en un proceso de *handover* entre dos eNBs, el cambio del plano de usuario puede únicamente derivar en un cambio del servicio portador S1 entre los eNBs implicados y el S-GW, manteniéndose sin cambios el resto del plano de usuario (camino entre S-GW y P-GW).
- La funcionalidad de punto de anclaje también se aplica a la gestión de movilidad con las otras redes de acceso 3GPP (UTRAN y GERAN). De esta forma, equipos de usuario que se conecten a la red LTE a través de UTRAN o GERAN, disponen también de un S-GW asociado en la red troncal EPC por el que fluye su plano de usuario. Esta configuración, junto con otras posibilidades de interconexión con otras redes que ofrece LTE, se describirán más detalladamente en el Capítulo 7.
- Almacenamiento temporal de los paquetes IP de los usuarios en caso de que los terminales se encuentren en modo *idle*. En la red LTE, el plano de usuario entre S-GW y el equipo de usuario puede “desactivarse” cuando no haya tráfico para transmitir. Es decir, aunque las conexiones y servicios portadores EPS permanezcan activos, un

terminal puede encontrarse en estado *idle* y, por tanto, no estar conectado a ningún eNB. Así pues, cuando se recibe tráfico de la red externa dirigido a un usuario en modo *idle*, este tráfico llega hasta la entidad S-GW a cargo de ese usuario, que retiene temporalmente los paquetes IP e inicia (a través de la señalización pertinente con la entidad MME) el reestablecimiento del plano de usuario hasta el equipo de usuario.

- Encaminamiento del tráfico de usuario. Como todo el tráfico de un usuario fluye a través de una pasarela S-GW, ésta alberga la información y funciones de encaminamiento necesarias para dirigir el tráfico de subida (tráfico IP proveniente de los equipos de usuario) hacia la pasarela (o pasarelas) P-GW que corresponda y el tráfico de bajada (proveniente de las pasarelas P-GW) hacia el eNB a través del cual se encuentra conectado el equipo de usuario. Es importante destacar que, aunque un usuario puede tener múltiples conexiones establecidas con diferentes pasarelas P-GW de forma simultánea, todo el tráfico atraviesa una única S-GW.

2.6.2.3 PDN Gateway (P-GW).

Esta entidad es la encargada de proporcionar conectividad entre la red LTE y las redes externas (denominadas como *Packet Data Network*, PDN, en las especificaciones 3GPP). Es decir, a través de la entidad P-GW, un usuario conectado al sistema LTE resulta "visible" en la red externa. Por tanto, los paquetes IP generados por el usuario se inyectan en la red externa a través de esta pasarela y, viceversa, todo el tráfico IP dirigido a un terminal LTE proveniente de la red externa va a ser encaminado hasta el P-GW. Tal como se verá más en detalle en el Capítulo 3, un usuario tiene asignada como mínimo una pasarela P-GW desde su registro en la red LTE. Entre las principales funciones de la pasarela P-GW podemos destacar:

- Aplicación de las reglas de uso de la red (*i.e.*, *policy control*) y control de tarificación a los servicios portadores que tenga establecidos el terminal. Estas funciones forman parte del marco PCC introducido previamente en el apartado 2.6.1 y que se analizará más detalladamente en el Capítulo 3.
- La asignación de la dirección IP de un terminal utilizada en una determinada red externa se realiza desde la pasarela P-GW correspondiente. La dirección puede ser una dirección IPv4, IPv6 o bien un par de direcciones (IPv4, IPv6). El mecanismo de asignación de la dirección se sustenta en la señalización propia de la red LTE (e.g., el terminal recibe la dirección IP a través de los protocolos NAS) o bien en la utilización de protocolos propios de redes IP como DHCP [19].
- La pasarela P-GW actúa de punto de anclaje para la gestión de movilidad entre LTE y redes no 3GPP. Tal como se verá en el Capítulo 7, la pasarela alberga funciones de Home Agent (HA) para proporcionar continuidad de servicio en caso de utilizar el protocolo Mobile IPv4 (MIPv4) [19] para gestionar la movilidad entre la red LTE y, por ejemplo, una red WiMAX. Además de MIPv4, la pasarela incluye soporte de movilidad para los protocolos Dual Stack MIPv6 (DSMIPv6) y Proxy MIPv6 (PMIPv6).

- El tráfico IP que transcurre por la pasarela P-GW es procesado a través de un conjunto de filtros que asocian cada paquete IP con el usuario y servicio portador EPS correspondiente. Esto permite, por un lado, aplicar las reglas de uso y tarificación antes comentadas, y por otro, aplicar funciones de inspección y verificación de la validez de los paquetes IP que cursa la red (*packet screening*). De esta forma, la pasarela puede descartar los paquetes IP que sean considerados como tráfico anómalo (e.g., un equipo de usuario envía paquetes con una dirección o puertos para los que no está autorizado).

2.6.2.4 HSS (Home Subscriber Server).

El HSS es la base de datos principal del sistema 3GPP que almacena la información de los usuarios de la red. La información contenida en el HSS abarca tanto información relativa a la suscripción del usuario (i.e., perfil de suscripción) como información necesaria para la propia operativa de la red. La base de datos HSS es consultada, y modificada, desde las diferentes entidades de red encargadas de proporcionar los servicios de conectividad o servicios finales (desde, e.g., MME de red troncal EPC, SGSN de la red GPRS, MSC del dominio de circuitos y también desde servidores de control del subsistema IMS). El HSS contiene tanto información permanente que sólo puede ser cambiada mediante procesos administrativos (e.g., campos creados al dar de alta a un usuario en la red o cambiar las condiciones de su contrato), así como información temporal que cambia a raíz de la propia operación del sistema (e.g., localización del terminal dentro de la zona de servicio del sistema). Así, entre la información almacenada en el HSS podemos destacar: identificadores universales del usuario (e.g., *International Mobile Subscriber Identity*, IMSI), identificadores de servicio (e.g., *Mobile Station ISDN*, MSISDN); información de seguridad y cifrado (vectores de autenticación); información de localización del usuario en la red (identificador de la entidad de control, e.g. MME, que proporciona el plano de control hacia un determinado usuario); e información necesaria para la provisión de los servicios de acuerdo con las condiciones establecidas en el contrato de suscripción (e.g., identificador de la red externa y parámetros de calidad de servicio del servicio portador por defecto).

La entidad HSS se estandarizó en 3GPP R5 en base a la integración de dos entidades definidas inicialmente en redes GSM y que se denominan HLR (Home Location Register) y AuC (Authentication Center), a las que se añadieron funciones adicionales necesarias para soportar el acceso y la operativa del sistema LTE. En la Release 8 correspondiente al sistema LTE, el HSS abarca:

- El subconjunto de funciones de las entidades HLR/AuC necesarias para el funcionamiento del dominio de paquetes EPC, así como GPRS. El acceso a HSS desde la red EPC se realiza desde la entidad de red MME mediante la interfaz S6a.
- El subconjunto de funciones de las entidades HLR/AuC necesarias para el funcionamiento del dominio CS.
- Funciones de soporte asociadas a las funciones de control del subsistema IMS como la gestión de información relativa a la suscripción de servicios IMS y el almacenamiento de perfiles de usuario asociados a servicios IMS.

La información almacenada en el HSS se detalla en la especificación 3GPP TS 23.008 [16]. Las entidades de red que acceden a la base de datos HSS para gestionar el acceso al servicio de conectividad de la red troncal EPC son las siguientes. Cuando el acceso se realiza a través de E-UTRAN, la entidad MME es la que interactúa con la base de datos a través de la interfaz S6a. Cuando el acceso es a través de UTRAN o GERAN, el acceso a HSS se realiza desde el Server GPRS Support Node (SGSN) mediante la interfaz S6d. Cuando el acceso es a través de redes no 3GPP, el acceso se canaliza a través del servidor AAA mediante la interfaz SWz. Los detalles de los accesos a través de redes diferentes a E-UTRAN se abordan en el Capítulo 7.

2.6.2.5 Interfaz P-GW ↔ Redes Externas (SGi)

A través de la interfaz SGi se realiza la interconexión de la pasarela P-GW de la red LTE con redes externas IP. La red externa puede ser tanto una red pública (e.g., Internet) como cualquier otra red privada (e.g., intranet corporativa, red de un ISP, red interna del propio operador para la provisión, por ejemplo, de servicios IMS). La interfaz SGi es equivalente a la interfaz Gi especificada para la interconexión de la pasarela GGSN del dominio GPRS con redes externas [20].

La interfaz SGi soporta la interconexión tanto a redes IPv4 como IPv6. Desde la perspectiva de la red externa, la pasarela P-GW es vista como un *router* IP convencional. Sobre esta base, existen dos modelos básicos de interconexión de la red LTE con la red externa: acceso transparente y acceso no transparente. Ambos modelos se ilustran en la Figura 2.15 y se describen a continuación.

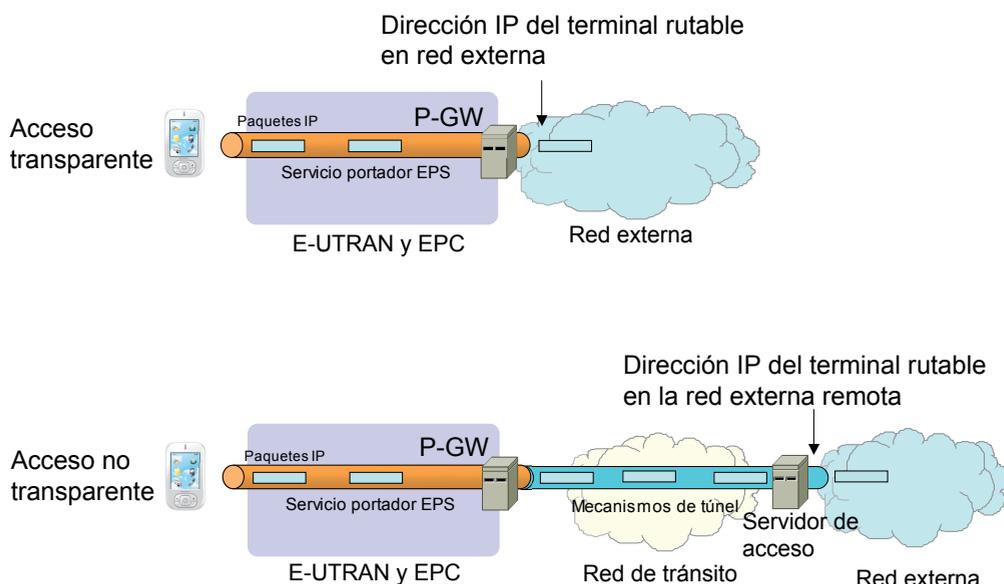


Figura 2.15 Tipos de interconexión a través de SGi

Bajo el modelo de interconexión transparente, la dirección IP asignada al terminal es válida en la propia interfaz SGi con la red externa, de forma que el terminal es "visible" en la red externa a la que proporciona acceso la pasarela P-GW a través de dicha dirección. En cambio, en el modelo no transparente, la red LTE ofrece un acceso a una red externa remota de forma que el espacio de direcciones utilizado por los terminales pertenece al espacio de direcciones de la red externa remota. La conexión entre la red LTE y la red remota admite diferentes soluciones tales como el establecimiento de un túnel IP (e.g., IPsec, GRE, etc.) entre la pasarela P-GW y el servidor de acceso remoto. En el modelo no transparente, puede ser necesario que la pasarela P-GW participe en, e.g., funciones de autenticación del usuario y asignación de direcciones dinámicas pertenecientes a la red remota.

La interfaz SGi y los diferentes tipos de acceso a la red externa se especifican en 3GPP TS 29.061 [17].

2.6.2.6 Interfaces P-GW ↔ S-GW (S5 y S8)

Las interfaces S5 y S8 proporcionan el soporte para la transferencia de paquetes de usuario entre las pasarelas S-GW y P-GW. La interfaz S5 se utiliza en situaciones donde ambas pasarelas pertenecen a la misma red mientras que la interfaz S8 es la utilizada en caso de escenarios de itinerancia (*roaming*) donde el S-GW pertenezca a la red visitada (denominada *Visited* PLMN) y el P-GW a la red matriz (denominada *Home* PLMN).

Ambas interfaces S5 y S8 admiten dos implementaciones diferentes: una basada en el protocolo GTP y otra basada en el protocolo PMIPv6. La implementación basada en GTP proporciona funciones de creación/eliminación/modificación/cambio del servicio portador del plano de usuario entre S-GW y P-GW de los usuarios conectados a la red LTE. En el caso de la implementación basada en PMIPv6, no se soporta ninguna gestión de servicios portadores entre P-GW y S-GW de forma que, estrictamente el servicio portador EPS de la red LTE se extiende desde el equipo de usuario hasta el S-GW (y no hasta la pasarela P-GW, como sería el caso de utilizar GTP en la interfaz S5/S8). Ello es debido a que el protocolo PMIPv6 está concebido para ofrecer exclusivamente un servicio de movilidad entre el S-GW y el P-GW y no dispone de los mecanismos necesarios para señalar parámetros de QoS. Por tanto, las interfaces S5 y S8, en el caso de la variante basada en PMIPv6, básicamente ofrecen un servicio de conectividad entre las pasarelas implicadas de forma que todos los flujos de datos son llevados a través del mismo túnel (sin distinción de servicios portadores).

La descripción funcional de la interfaz S5/S8 basada en GTP se aborda en TS 23.401 [14] y el protocolo en TS 29.274 [21]. La funcionalidad de la variante PMIPv6 se proporciona en TS 23.402 [15] y el protocolo en TS 23.275 [24].

2.6.2.7 Interfaz MME ↔ S-GW (S11)

Esta interfaz permite controlar la operativa del plano de usuario en la red troncal EPC des-

de la entidad de red MME. Así, los procedimientos soportados en esta interfaz permiten la creación/eliminación/modificación/cambio de los servicios portadores que los terminales tienen establecidos a través de la red troncal LTE. En este sentido, dado que la entidad MME es la entidad de control que termina los protocolos NAS con los usuarios, la interfaz S11 permite establecer el nexo del plano de control con las funciones del plano de usuario de la red troncal LTE.

Las acciones a realizar sobre la interfaz S11 tienen su origen en los diferentes eventos relacionados con la señalización entre el terminal y la red troncal (e.g., registro de un terminal en la red LTE, incorporación de un nuevo servicio portador EPS a una conexión, establecimiento de una nueva conexión con otra red externa, indicación de *handover*, etc.) así como en eventos originados desde la propia red troncal EPC (e.g., notificación de que existe tráfico almacenado en el S-GW pendiente de ser enviado a un terminal en modo *idle*, modificación de servicios portadores iniciada por la pasarela P-GW, etc.).

La interfaz S11 también da soporte al proceso de reubicación de la pasarela S-GW asociada a un terminal mediante la transferencia de contextos entre la pasarela antigua y la nueva (en el contexto se almacena toda la información relacionada con configuración del plano de usuario en el S-GW). Asimismo, durante la ejecución de un procedimiento de *handover* que implique una reubicación de S-GW, el nodo MME controla los mecanismos de transferencia de paquetes entre las dos posibles pasarelas involucradas.

La funcionalidad de esta interfaz se recoge en 3GPP TS 23.401 [5] y el protocolo GTPv2-C que da soporte a esta interfaz se especifica en TS 29.274 [21].

2.6.2.8 Interfaz MME ↔ MME (S10)

La interfaz S10 se define entre dos entidades MME. Su principal función es el soporte del mecanismo de reubicación de la entidad MME. De esta forma, cuando la entidad MME que controla a un determinado equipo de usuario debe cambiarse (debido, por ejemplo, a su movilidad), a través de la interfaz S10 se realiza la transferencia del contexto de dicho usuario entre MMEs. En este caso, el contexto de un terminal en una entidad MME abarca toda la información relacionada con la gestión de los servicios portadores (e.g., configuración de las conexiones activas) y la gestión de movilidad (e.g., área de seguimiento donde se encuentra localizado el terminal), junto con otra información de seguridad (claves de cifrado), características del terminal (*classmark*), parámetros de suscripción del usuario obtenidos desde el HSS, etc.

La funcionalidad de esta interfaz se recoge en 3GPP TS 23.401 [5] y el protocolo GTPv2-C que da soporte a esta interfaz se especifica en TS 29.274 [21].

2.6.2.9 Interfaz HSS ↔ MME (S6a)

Esta interfaz permite la transferencia de información entre la base de datos HSS y la enti-

dad del plano de control MME de la red troncal EPC. A través de la interfaz S6a se da soporte a diferentes funciones:

- Mantenimiento de información de gestión de la localización. La base de datos HSS mantiene unos campos que contienen la identificación del nodo MME que controla a cada usuario registrado en la red. Esta información la actualiza el nodo MME correspondiente a través de la interfaz S6a. Esta información permite que cuando un terminal se conecta a un MME nuevo, dicho MME pueda recuperar información relativa al nodo MME que previamente dió servicio al terminal de cara a realizar la reubicación pertinente.
- Autorización de acceso a la red LTE. La base de datos HSS almacena los datos de suscripción de los usuarios que condicionan el acceso a los servicios que ofrece la red. El perfil de suscripción de un usuario se transfiere desde el HSS al nodo MME, que es la entidad encargada de ejecutar las comprobaciones pertinentes. Así, a modo de ejemplo, la entidad MME puede autorizar o no la conexión de un usuario a una red externa en función de que dicha red externa forme parte de lista de redes externas permitidas según la suscripción del usuario.
- Autenticación de los usuarios. A través de S6a, el nodo MME se descarga la información que permite llevar a cabo el procedimiento de autenticación desde el MME. Esta información se denomina vector de autenticación EPS. En el apartado 3.4 del Capítulo 3 se proporcionan más detalles de la función de autenticación.
- Notificación y descarga de la identidad de la pasarela P-GW que utiliza un usuario en una conexión. El almacenamiento en la base de datos HSS de información relativa a las pasarelas P-GW que dan servicio al usuario se utiliza para proporcionar soporte a los mecanismos de movilidad entre LTE y otras redes no 3GPP.

La interfaz también soporta escenarios de itinerancia donde una entidad MME de la red de un operador puede acceder a la base de datos HSS de otro operador. En el apartado 2.6.5 se describen diferentes escenarios de itinerancia.

La interfaz S6a se basa en el protocolo Diameter. La funcionalidad de la interfaz se recoge en TS 23.401 [5] y la especificación de la extensión (i.e., *Diameter application*) del protocolo se aborda en TS 29.272 [25].

2.6.3 Protocolos

En este apartado se describen las torres de protocolos utilizadas en las interfaces descritas en el apartado anterior. Dado que muchas de las interfaces están basadas en una misma torre de protocolos, la descripción se realiza en base a los diferentes protocolos utilizados.

Los protocolos NAS (*Non Access Stratum*) entre el equipo de usuario y la entidad de red MME también se detallan en este apartado.

Es importante destacar que en la red troncal EPC todas las interfaces se soportan sobre una torre de protocolos que utiliza IP como capa de red, a diferencia de lo que ocurría con las familias anteriores de sistemas 3GPP donde algunas interfaces estaban soportadas sobre los protocolos ATM o SS7 [20].

2.6.3.1 Interfaces basadas en GTP-U

Todas las interfaces para el transporte de información de plano de usuario entre los diferentes elementos de la red troncal EPC se soportan a través del protocolo GTP-U, excepto la variante de la interfaz S5/S8 basada en PMIPv6. La torre de protocolos utilizada en las interfaces basadas en GTP-U y el listado de dichas interfaces se proporciona en la Figura 2.16. Nótese que en la tabla ilustrada en la figura se indica el uso de GTP-U también sobre las interfaces S4 y S12 que forman parte de la solución de *interworking* entre EPC y el resto de redes 3GPP y que serán analizadas en más detalle en el Capítulo 7. Además, el protocolo GTP-U también se utiliza en el plano de usuario de las interfaces internas de E-UTRAN, S1-U y X2-U, tal como se ha visto en el apartado 2.5.

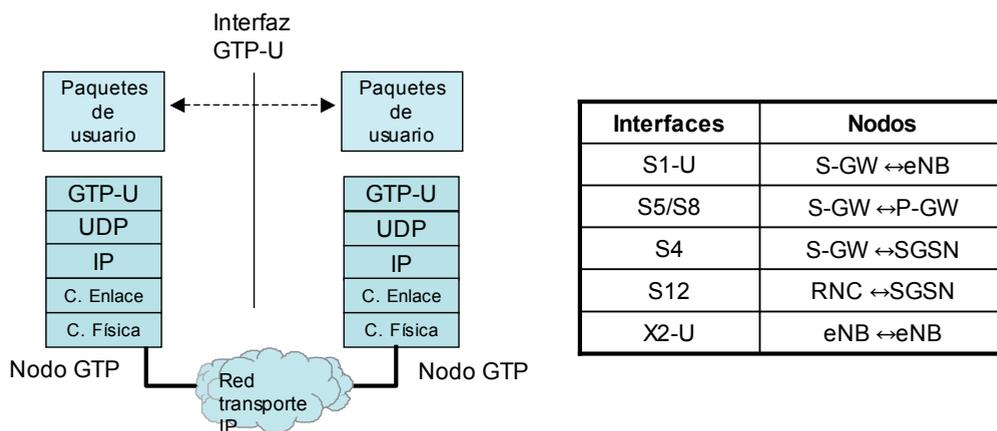


Figura 2.16 Interfaces basadas en GTP-U

El protocolo GTP-U fue desarrollado por 3GPP para dar respuesta a la implementación del servicio GPRS. En este sentido, el plano de usuario entre los nodos de red del dominio GPRS así como el plano de usuario de la interfaz lu-PS de UTRAN se soportan también sobre dicho protocolo.

GTP-U proporciona un mecanismo de encapsulado para el envío de paquetes de usuario (e.g., paquetes IP del usuario) entre nodos de una red IP. Los paquetes que corresponden a un mismo servicio portador EPS se transportan con un identificador de túnel único denominado

TEID (Tunnel Endpoint Identifier). A modo ilustrativo, en la Figura 2.17 se representa la implementación de un túnel entre las pasarelas S-GW y P-GW (interfaz S5/S8) mediante GTP-U. Tal como se observa en la figura, los paquetes IP del usuario llegan a la pasarela S-GW provenientes desde el equipo de usuario a través de los servicios portadores radio y S1. Las direcciones IP origen y destino de los paquetes de usuario recibidos en el S-GW contienen, respectivamente, la dirección asignada al terminal móvil y la dirección del equipo de la red externa al que vaya dirigido el paquete. Nótese que estas direcciones no tienen porque pertenecer al espacio de direcciones IP utilizado en la red de transporte que une las pasarelas S-GW y P-GW, de ahí la necesidad de establecer el túnel. Así, para proceder al envío de estos paquetes IP de usuario hacia la pasarela P-GW, el nodo S-GW los encapsula mediante el protocolo GTP-U. La cabecera del protocolo GTP-U ocupa un mínimo de 6 bytes y contiene el identificador de túnel TEID, junto con otros parámetros tales como identificadores de secuencia y longitud del paquete. El paquete GTP resultante tiene como dirección IP origen la dirección de la pasarela S-GW y como dirección destino la dirección IP de la pasarela P-GW. De esta forma, el paquete GTP puede ser encaminado en la red de transporte IP que une a ambas pasarelas. Una vez el paquete GTP llega a la pasarela P-GW, ésta extrae el paquete IP del usuario y lo inyecta en la red externa.

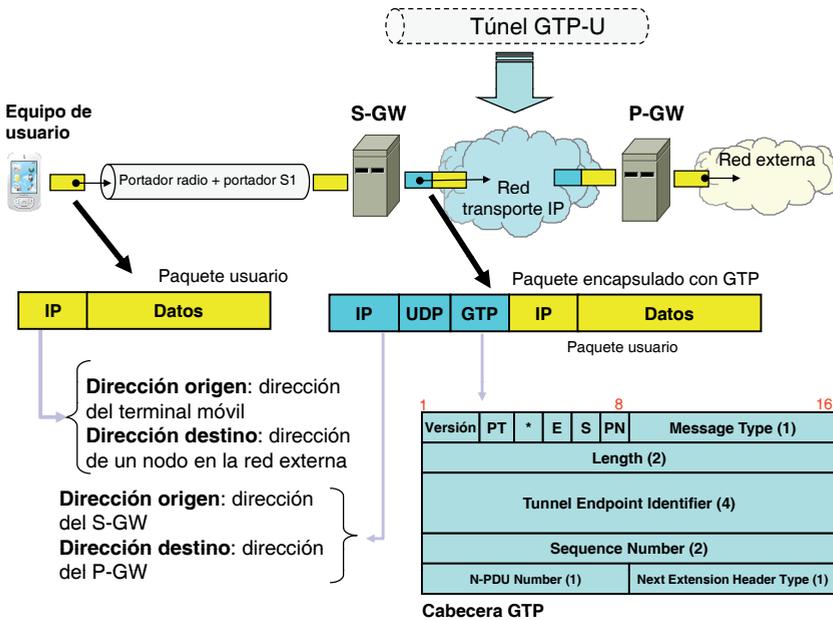


Figura 2.17 Ilustración del funcionamiento de un túnel GTP-U

El establecimiento de un túnel GTP-U consiste básicamente en la elección del identificador TEID asociado a un determinado servicio portador EPS en ambos extremos del túnel. La señalización necesaria para establecer el túnel se realiza mediante otros protocolos como GTP-C o S1-MME.

El protocolo GTP-U, y en particular su versión GTPv1-U, se utiliza tanto en LTE como en UMTS y se especifica en TS 29.281 [26].

2.6.3.2 Interfaces basadas en GTP-C

El protocolo GTP-C soporta un conjunto de funciones que pueden clasificarse en torno a los siguientes aspectos:

- Gestión de sesiones. A través de los mensajes y procedimientos de señalización especificados para GTP-C, la red gestiona la creación de túneles GTP-U entre las entidades de la red por donde transcorre el plano de usuario. Dichos túneles forman parte de la propia gestión de sesiones en la red, mediante el establecimiento, mantenimiento/actualización y liberación de conexiones PDN y servicios portadores EPS. Los fundamentos y procedimientos asociados a la gestión de sesiones en el sistema LTE se abordan más detalladamente en el apartado 3.2 del Capítulo 3.
- Gestión de movilidad. Mediante el protocolo GTP-C se llevan a cabo algunos de los procedimientos asociados con la gestión de movilidad tales como la transferencia de los contextos de información de los usuarios entre las entidades de red en casos de reubicación de las mismas.

En la Figura 2.18 se ilustra la torre de protocolos de las interfaces basadas en GTP-C y la relación de las interfaces. El protocolo GTP-C se utiliza en las interfaces S3, S4, S5/S8, S10, S11 y S16. Nótese que no hay una correspondencia directa entre los interfaces que utilizan GTP-U en el plano de usuario (véase la Figura 2.16) y GTP-C en el plano de control, ya que, tal como se ha comentado, el protocolo GTP-C abarca otras funciones además de la gestión de túneles GTP-U.

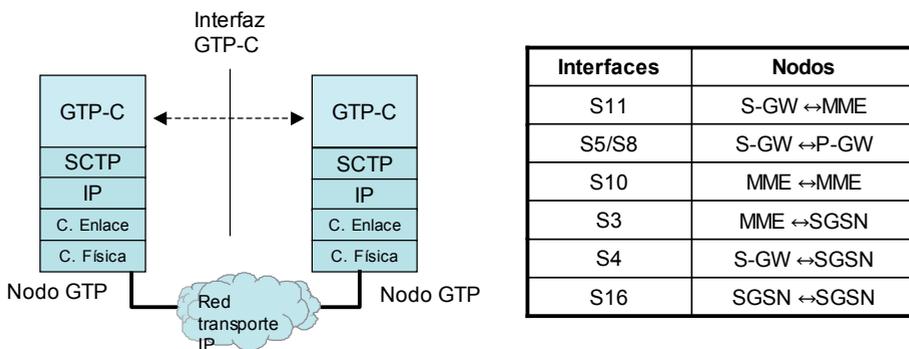


Figura 2.18 Interfaces basadas en GTP-C

La versión del protocolo GTP-C utilizada en las interfaces de LTE, denominada como GTPv2-C, se especifica en TS 29.274 [21]. En cambio, la versión del protocolo GTP utilizada en interfaces propias del sistema UMTS (e.g., interfaces entre los elementos de la red tron-

cal GPRS) se cubre en el documento TS 29.060 [22]. Detalles adicionales de la operativa del protocolo GTP-C se analizan en el Capítulo 3 durante la descripción de los procedimientos de gestión de sesiones y de gestión de movilidad.

2.6.3.3 Interfaces basadas en Diameter

El protocolo Diameter es una evolución del protocolo RADIUS, inicialmente concebido para sustentar funciones de *Autenticación, Autorización y Accounting* (AAA). Diameter mejora las prestaciones de su antecesor RADIUS en aspectos tales como seguridad, robustez a pérdidas de mensajes, así como en su extensibilidad que permite el uso del protocolo para aplicaciones fuera del ámbito de AAA.

El protocolo Diameter se utiliza en un elevado número de interfaces del nuevo sistema LTE. En la Figura 2.19 se ilustra la torre de protocolos sobre la que se sustenta Diameter junto con una tabla donde se indican todas las interfaces del sistema LTE basadas en dicho protocolo. La transferencia de los mensajes Diameter entre nodos se realiza a través de un protocolo de transporte orientado a conexión como TCP o SCTP.

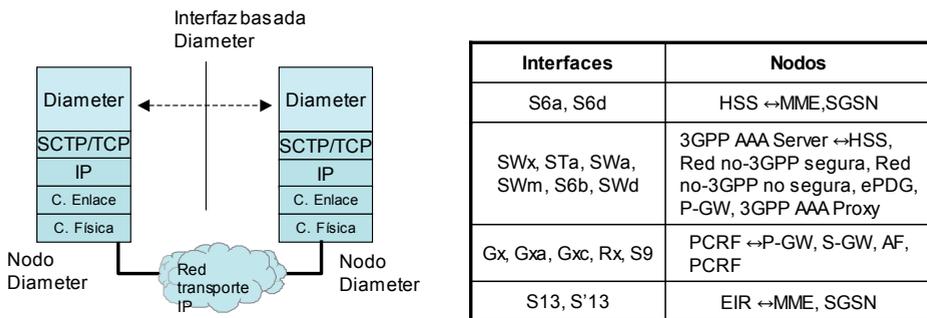


Figura 2.19 Interfaces basadas en Diameter

El protocolo Diameter se estructura en torno a una protocolo de base (*Diameter base standard* definido en RFC 3588 [32]) y un número de extensiones denominadas aplicaciones. El protocolo de base aporta las funcionalidades comunes del protocolo: formatos de los mensajes y elementos de información genéricos (e.g., *Attribute Value Pairs*, AVPs), mecanismos de transferencia de mensajes, descubrimiento de capacidades de las entidades Diameter, aspectos de seguridad, etc. Las "aplicaciones" definen los mensajes adicionales y los procedimientos necesarios para adaptar el uso de Diameter al soporte de una determinada funcionalidad. Entre las aplicaciones de Diameter más relevantes estandarizadas por IETF se encuentran: *Network Access Server Application* (aplicación de Diameter para servicios AAA en el marco de control de acceso a redes, definido en la RFC 4005 [33]) y *Credit Control Application* (aplicación de Diameter para la implementación de sistemas de tarificación on-line, como sistemas de pre-pago, definido en la RFC 4006 [34]). Además de IETF, otros organismos

también pueden llevar a cabo la especificación de nuevas aplicaciones del protocolo, como es el caso de 3GPP. Estas aplicaciones de Diameter se denominan como “*vendor-specific*” y se les asigna un identificador de aplicación a través de IANA (Internet Assigned Numbers Authority⁵). Así pues, 3GPP ha definido varias aplicaciones “*vendor-specific*” para la implementación de diferentes interfaces del sistema mediante extensiones del protocolo Diameter. Cada una de estas aplicaciones de Diameter se recoge en un documento de especificación técnica del 3GPP. A modo de ejemplo, la aplicación de Diameter para la interfaz S6a/S6d se define en 3GPP TS 29.272 [25].

2.6.3.4 Interfaces basadas en PMIPv6

El protocolo PMIPv6 (Proxy MIPv6) es un protocolo especificado por IETF en la RFC 5213 [35] para gestionar la movilidad a nivel de capa de red IP. El protocolo PMIPv6 ha sido adoptado por 3GPP para su posible utilización en la interfaz S5/S8 entre las pasarelas S-GW y P-GW, como alternativa al uso del protocolo GTP especificado por 3GPP. Al igual que la alternativa basada en GTP, PMIPv6 resuelve la movilidad de forma transparente al equipo de usuario, es decir, sin necesidad de que éste participe en la señalización pertinente. Este modelo de gestión de movilidad se conoce como gestión de movilidad “*network-based*”, en contraposición al modelo “*host-based*” establecido por el protocolo MIP donde los nodos extremos (equipos de usuarios) participan en la gestión de movilidad [19]. En la Figura 2.20 se ilustra el ámbito de utilización del protocolo de movilidad PMIPv6 junto con sus componentes funcionales. El protocolo define una entidad LMA (*Local Mobility Anchor*) que realiza funciones similares a un *Home Agent* (HA) en MIP. Fundamentalmente el LMA mantiene una asociación entre la dirección IP que tiene asignada el terminal (y que no pertenece al espacio de direcciones IP de la red de transporte IP que conecta el LMA y los MAGs) y la dirección IP hacia la que debe enviar los paquetes del usuario mediante un mecanismo de encapsulado. La dirección IP de envío de los paquetes es la dirección del *router* que alberga la funcionalidad de MAG (*Mobile Access Gateway*). De esta forma, en el caso de la interfaz S5/S8, todos los paquetes IP que llegan a la pasarela P-GW (LMA) desde la red externa y que contienen como dirección destino la dirección IP asignada a un terminal, son encapsulados y enviados mediante un túnel PMIPv6 a la pasarela S-GW correspondiente (MAG), y viceversa. Nótese que la transferencia de los paquetes IP entre las pasarelas S-GW y los equipos de usuario ya no compete a la operativa del protocolo y se aborda mediante los mecanismos soportados a tal efecto en las interfaces S1 y radio (servicios portadores radio y S1).

⁵ IANA es un órgano que depende de ICANN (Internet Corporation for Assigned Names and Numbers, <http://www.icann.org/>) que se encarga de la asignación y mantenimiento de identificadores únicos y sistemas de numeración necesarios para el funcionamiento de los protocolos utilizados en Internet.

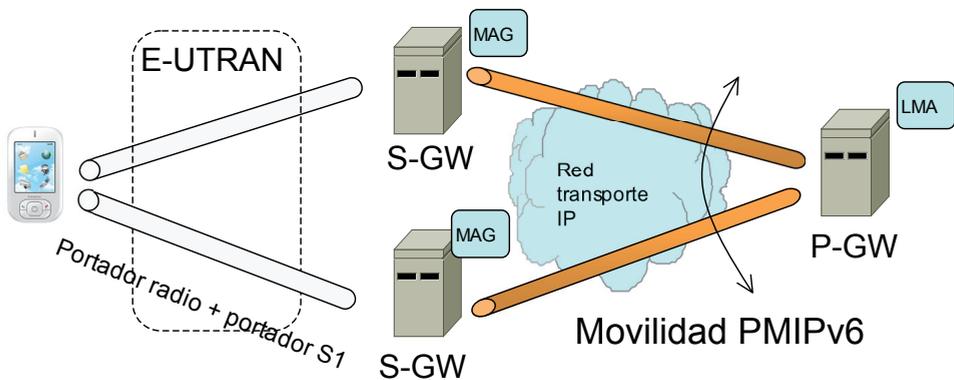
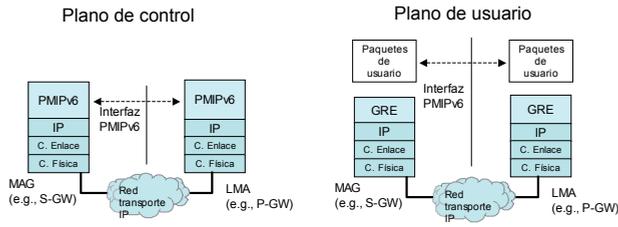


Figura 2.20 Ámbito y componentes del protocolo PMIPv6

En la Figura 2.21 se ilustran los planos de control y de usuario del protocolo PMIPv6, junto con las interfaces donde puede utilizarse el protocolo. El plano de control consiste en unos mensajes de señalización especificados en el protocolo que se envían en la parte de datos de los paquetes IP intercambiados entre MAGs y LMA. A través de dichos mensajes de señalización se controlan las asociaciones de direcciones y los túneles necesarios. El plano de usuario del protocolo se basa en el establecimiento de un túnel que permite enviar de forma transparente los paquetes IP de los usuarios (con direcciones origen y destino pertenecientes al espacio de direcciones de la red externa) entre MAGs y LMA (que pueden utilizar un espacio de direcciones diferente a las de la red externa). Para ello, los paquetes IP de usuario se encapsulan dentro de la carga útil de paquetes IP mediante el protocolo GRE (*Generic Routing Encapsulation*, RFC 2784 [36]). El protocolo GRE añade unas cabeceras al paquete IP de usuario que permiten asociar cada paquete con la conexión PDN a la que pertenece. En cambio, a diferencia de GTP, los túneles GRE utilizados por PMIPv6 no distinguen entre el tráfico que pertenece a diferentes servicios portadores EPS establecidos en el contexto de una misma conexión PDN (i.e., los identificadores de túnel GRE hacen referencia a una conexión PDN mientras que los identificadores de túnel GTP-U hacen referencia a un servicio portador EPS). Tal como se verá en el apartado 3.2 del Capítulo 3, esta diferencia hace que los servicios portadores EPS se extiendan entre la pasarela S-GW y el equipo de usuario cuando se utiliza PMIPv6 en la interfaz S5/S8. Por el contrario, mediante el uso de GTP en S5/S8 los servicios portadores EPS se extienden desde la pasarela P-GW hasta el equipo de usuario.



Interfaces	Nodos
S5/S8	S-GW ↔ P-GW
S2a	P-GW ↔ Red no 3GPP trusted
S2b	P-GW ↔ ePDG

Figura 2.21 Interfaces basadas en PMIPv6

Además de la interfaz S5/S8, el protocolo PMIPv6 es uno de los protocolos especificados para el soporte de movilidad entre LTE y redes no 3GPP. En particular, las interfaces S2a y S2b están basadas en PMIPv6. En el Capítulo 7 se aportan más detalles de estas dos interfaces.

2.6.3.5 Protocolos NAS

Los protocolos NAS son los protocolos desarrollados por el 3GPP para llevar a cabo la gestión de movilidad de los equipos de usuario (*EPS Mobility Management, EPM*) y la gestión de las sesiones para el establecimiento de la conectividad entre el equipo de usuario y la pasarela P-GW (*EPS Session Management, ESM*). Los protocolos NAS se soportan entre el equipo de usuario y un nodo MME y se han desarrollado específicamente para E-UTRAN, aunque se mantienen muchas similitudes con los protocolos NAS utilizados en UMTS [20] (e.g., Session Management, SM, y GPRS Mobility Management, GMM, del dominio GPRS). En la Figura 2.22 se ilustra el alcance de los protocolos NAS en la red LTE.

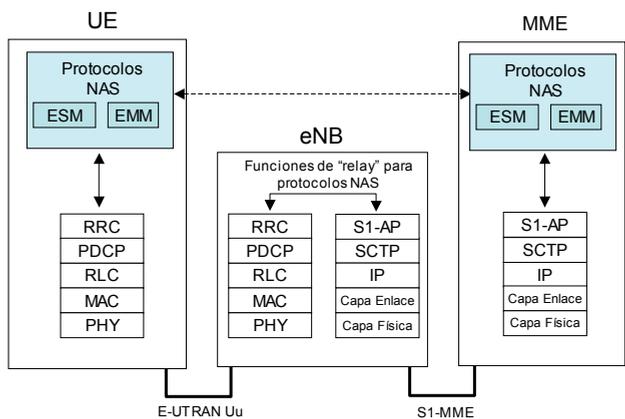


Figura 2.22 Protocolos NAS entre UE y MME

Los protocolos ESM y EMM se especifican en 3GPP TS 24.301 [37].

2.6.3.5.1 Protocolo NAS para la gestión de movilidad (EMM)

El protocolo EMM proporciona los procedimientos necesarios para el control de la movilidad de un equipo de usuario que utiliza E-UTRAN para el acceso a la red troncal EPC.

En particular, entre los procedimientos soportados por el protocolo EMM se encuentran los mecanismos de “registro” y “cancelación de registro” del usuario en la red LTE (procedimientos denominados como *Network Attach* y *Detach* en las especificaciones) y la actualización del área de seguimiento (procedimiento denominado como *Tracking Area Update*). Mediante estos procedimientos se gestiona la accesibilidad a los servicios de la red LTE de los usuarios (la realización del registro en la red LTE es necesaria para que el usuario pueda iniciar o ser contactado para proceder a la activación de un servicio).

En el caso de terminales que se encuentren en estado *idle*, mediante el protocolo EMM se soporta el procedimiento de aviso (*paging*). En particular, el mensaje de aviso es un mensaje de señalización generado por el protocolo EMM que se distribuye a los terminales mediante las funciones disponibles en la interfaz S1-MME. El procedimiento de aviso lo utiliza la red troncal EPC para forzar el re-establecimiento de la señalización de control con un equipo de usuario que se encuentre en modo *idle*. Asimismo, el protocolo EMM soporta un procedimiento de petición de servicio (denominado *Service Request*) por parte del equipo de usuario cuyo propósito es permitir “reactivar” el plano de usuario entre el S-GW y un equipo de usuario que se encuentre en modo *idle*. La petición del servicio la realiza el equipo de usuario cuando, e.g., tiene paquetes IP pendientes de ser transmitidos.

El protocolo EMM también contempla procedimientos que permiten a la red interrogar al equipo terminal para el envío de identificadores tales como el IMSI (*International Mobile Subscriber Identity*) o el IMEI (*International Mobile Equipment Identity*) y llevar a cabo la autenticación del usuario (procedimiento denominado como *EPS Authentication and Key Agreement*, AKA). En particular el procedimiento EPS AKA permite la autenticación mutua entre usuario y red LTE así como el establecimiento de una clave maestra a partir de la cual se derivan las claves de cifrado e integridad.

Finalmente, también cabe destacar que es posible llevar a cabo el envío de información diversa entre el equipo de usuario y la red troncal EPC mediante un procedimiento de transporte sobre mensajes NAS soportado por el protocolo EMM. A través de dicho procedimiento se puede soportar, por ejemplo, la transferencia de mensajes SMS a través de la red LTE. Los mensajes SMS se envían encapsulados en mensajes NAS EMM.

En el Capítulo 3 se analizan más detalladamente los procedimientos más relevantes relacionados con la gestión de movilidad y de seguridad.

2.6.3.5.2 Protocolo NAS para la gestión de las sesiones (ESM)

El protocolo ESM sustenta los procedimientos necesarios entre el equipo de usuario y la red LTE para la gestión de los servicios portadores EPS cuando el equipo de usuario utiliza E-UTRAN.

Entre los procedimientos soportados por el protocolo ESM se encuentran los procedimientos de gestión (activación/desactivación/modificación) de los servicios portadores EPS. Tal como se verá en el Capítulo 3, además del servicio portador por defecto, pueden establecerse múltiples servicios portadores EPS dedicados que permiten aplicar un trato de QoS específico a un determinado flujo de paquetes. Estos procedimientos pueden realizarse en cualquier instante de tiempo, una vez el terminal se encuentra registrado y tiene establecido el servicio portador por defecto. En cuanto a la activación del servicio portador por defecto, una característica importante de LTE que no se contempla en UMTS y GPRS, es que su activación puede realizarse de forma conjunta con el procedimiento de registro, reduciéndose por tanto la señalización necesaria. Durante la activación del servicio portador por defecto, mediante el protocolo ESM se puede llevar a cabo la asignación de la dirección IP al equipo de usuario, aunque es importante señalar que LTE también soporta la asignación de la dirección a través de protocolos IETF en lugar de utilizar la señalización NAS.

El protocolo ESM también contempla un procedimiento que permite que un usuario solicite a la red el establecimiento de la conexión a una red externa. En respuesta a dicha petición, la red LTE puede proceder a activar el servicio portador por defecto con dicha red externa. El protocolo soporta también un mecanismo de petición de asignación de recursos (denominado como UE *requested bearer resource allocation procedure*). Este mecanismo permite que el equipo de usuario pueda notificar su necesidad de disponer de recursos que le permitan transferir un flujo de datos con unas determinadas características de QoS. Como respuesta a esta petición, la red LTE puede iniciar el establecimiento de un servicio portador EPS dedicado. Tal como se verá en el apartado 3.2 del Capítulo 3, en el sistema LTE, a diferencia de las redes UMTS, el establecimiento de servicios portadores se controla siempre desde la red troncal. Por tanto, este mecanismo permite disponer de una alternativa para que el terminal pueda solicitar el inicio de la activación del servicio portador, otorgando más flexibilidad para el soporte de aplicaciones cuya señalización no se controle directamente a través de plataformas de servicios.

En el Capítulo 3 se analizan más detalladamente los procedimientos más relevantes relacionados con la gestión de las sesiones en la red LTE.

2.6.4 Configuraciones de la red EPC

Las entidades de red en base a las que se describe la arquitectura de los sistemas LTE son entidades funcionales. Así, una entidad de red en 3GPP se concibe como una entidad "lógica" que cubre una funcionalidad perfectamente delimitada. Por tanto, una implementación concreta de la red LTE admite que diferentes entidades funcionales pueden residir en el mismo equipo físico.

En la Figura 2.23 se ilustran cuatro posibles implementaciones de la red troncal EPC en base a la ubicación física de las tres principales entidades de red que la componen: MME, S-GW y P-GW. Tal como se muestra en la opción (A), una implementación posible de la red troncal EPC consiste en integrar las tres entidades funcionales en un único equipo de red. Esta opción conduce a que el número de "saltos" o puntos de procesamiento del plano de usuario en la red LTE sea únicamente de dos (eNB y equipo de la red troncal S-GW+P-GW), con la consiguiente mejora en términos de latencia. No obstante, esta configuración no permite dimensionar por separado los recursos necesarios para soportar el plano de control y el de usuario de forma que el número de equipos de red troncal necesarios debe contemplar el peor de los casos. Nótese que, como el dimensionado de los recursos del plano de control depende principalmente del número de usuarios mientras que el dimensionado del plano de usuario está asociado al volumen de tráfico, la proporción entre el número de recursos necesarios para soportar ambos planos puede abarcar un amplio rango de valores atendiendo a la relación y evolución del número de usuarios y del volumen de tráfico que genera cada usuario en la red. De la misma manera, el montante de recursos necesarios para soportar la funcionalidad de punto de anclaje del plano de usuario (i.e., S-GW) y la funcionalidad de pasarela con redes externas (i.e., P-GW) tampoco guarda una relación de proporcionalidad clara. El número de equipos de red que alberguen la funcionalidad de P-GW puede depender en gran medida del número y tipología de las redes externas a las que se debe proporcionarse servicio mientras que el número de equipos S-GW está vinculado más directamente al número de usuarios y distribución geográfica de la red de acceso. Por tanto, implementaciones de equipos que alberguen ambas funciones, tal como sería el caso de la opción (A), limitan claramente la versatilidad y escalabilidad del sistema.

En el otro extremo, la opción (D) consiste en alojar cada entidad de red en un equipo físico diferente. Esta opción permite dimensionar más adecuadamente los recursos necesarios para cada funcionalidad: plano de control, plano de usuario y puntos de anclaje/pasarelas con redes externas. En este caso, el número de "saltos" del plano de usuario en la red es de tres (estación base, S-GW y P-GW).

Además de los dos casos extremos, la Figura 2.23 también muestra las otras dos posibilidades: la opción (B) permitiría explotar la separación de plano de control y usuario, y la opción (C) explotaría la separación de las funciones de anclaje y las funciones de interconexión a redes externas.

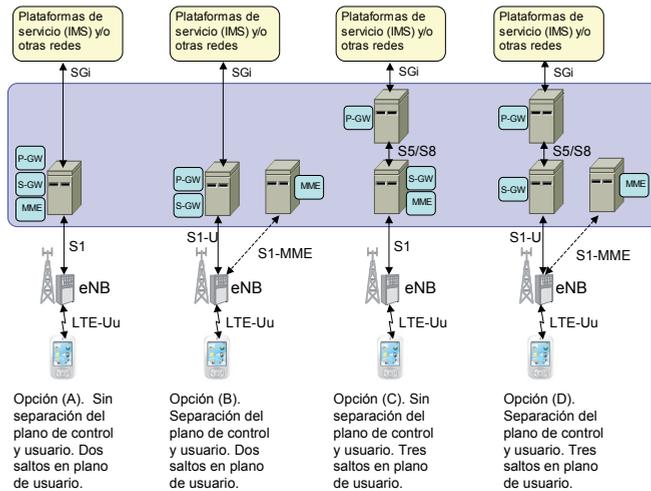


Figura 2.23 Posibles configuraciones de la red EPC

2.6.5 Soporte de itinerancia (roaming)

Una funcionalidad importante de las redes de comunicaciones móviles es el soporte del servicio de itinerancia (*roaming*). Este servicio permite que los usuarios puedan acceder a sus servicios de telecomunicación a través de las redes de otros operadores con los que no tienen establecida ninguna relación contractual (suscripción). A efectos de nomenclatura, el operador con el que el usuario tiene establecida la relación contractual para la prestación de servicios se conoce como operador matriz, y por extensión, la red de dicho operador constituye la red matriz (*Home Network*). La red de otro operador a la que el usuario puede tener acceso se denomina red visitada (*Visited Network*).

El sistema LTE especifica tres posibles configuraciones para la implementación de un servicio de itinerancia. Las diferentes configuraciones dependen de qué pasarela P-GW se utiliza para encaminar el tráfico con la red externa y de la capacidad de proporcionar acceso a los servicios propios del operador matriz. Las tres configuraciones son:

- I. Encaminamiento de tráfico a través de la red matriz, con acceso a los servicios de la red matriz.
- II. Encaminamiento de tráfico a través de la red visitada, con acceso a los servicios de la red visitada.
- III. Encaminamiento de tráfico a través de la red visitada, con acceso a los servicios de la red matriz.

La configuración de la arquitectura de red correspondiente a la opción (I) se ilustra en la Figura 2.24. En este caso, el nodo MME en la red visitada responsable de la terminación del plano de control con el terminal (protocolos NAS) accede a la base de datos HSS de la red

matriz (a través de la interfaz S6a) para poder obtener la información necesaria del usuario en itinerancia. De esta forma, el usuario en itinerancia puede autenticarse en la red visitada a partir de las credenciales que le han sido otorgadas por su operador matriz. Por otro lado, el establecimiento de los servicios de conectividad se realiza mediante la utilización de pasarelas P-GW del operador matriz. Para ello, se establece un túnel (a través de la interfaz S8) entre la pasarela S-GW que actúa de punto de anclaje en la red visitada y la pasarela P-GW que proporciona la interconexión con la red externa en la red matriz. Esta configuración permite que el usuario acceda a todos sus servicios como si estuviera conectado a través de la red de acceso de su operador matriz.

La Figura 2.25 representa la configuración de la red en el caso de la opción (II). Esta configuración permite que el tráfico generado por los usuarios en itinerancia se curse de forma local en las redes visitadas. Así, el acceso a las redes externas y/o plataformas de servicio se realiza mediante pasarelas P-GW pertenecientes a la red visitada. En cualquier caso, nótese como, al igual que en la opción (I), el nodo MME de la red visitada accede directamente a la base de datos HSS de la red matriz para obtener la información relativa al usuario en itinerancia. Asimismo, el control de las reglas de uso de la red y de tarificación, atendiendo a que cada operador puede establecer sus propias estrategias comerciales, puede realizarse en base al acceso al sistema PCC de la red matriz. De esta forma, las reglas de uso (e.g., parámetros de QoS) que se aplicarían en la red visitada vendrán determinadas por el operador matriz. Las reglas de uso pueden transferirse desde la red matriz de un usuario a la red visitada donde está recibiendo servicio a través de la interfaz S9 diseñada a tal efecto. La descripción del sistema PCC se aborda más detalladamente en el Capítulo 3.

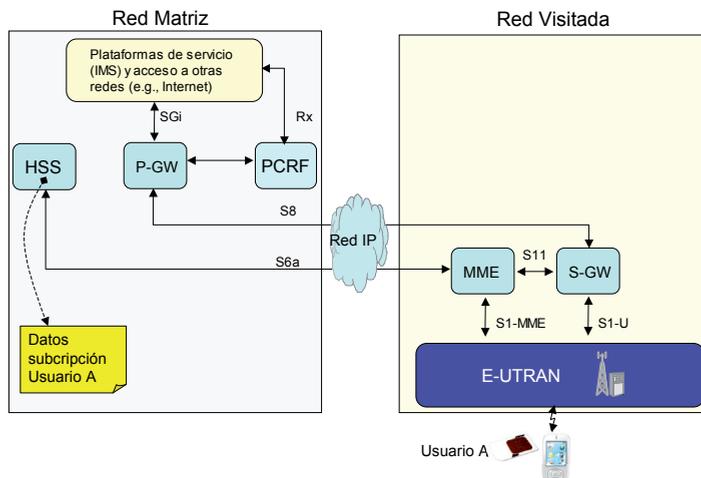


Figura 2.24 Encaminamiento de tráfico a través de la red matriz

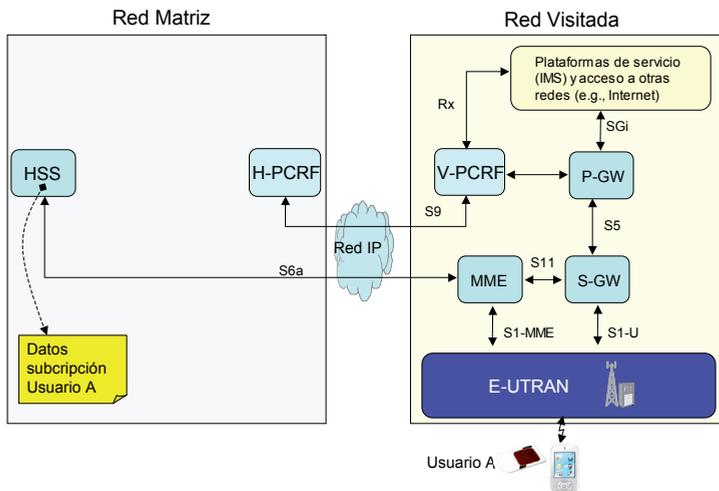


Figura 2.25 Encaminamiento de tráfico a través de la red visitada, con acceso a los servicios de la red visitada

Finalmente, la configuración correspondiente a la opción (III) se muestra en la Figura 2.26. Esta configuración resulta útil en el caso de que se pretenda que el usuario en itinerancia tenga acceso a las plataformas de servicio de su operador (el control de los servicios se realiza a través del operador matriz) pero quiera evitarse el encaminamiento de todo el tráfico a través de la red matriz. De esta forma, el tráfico de usuario que no necesariamente tenga que atravesar la plataforma de servicios puede acceder a las redes externas (e.g., Internet) sin necesidad de que tenga que ser transferido desde la red visitada a la red matriz.

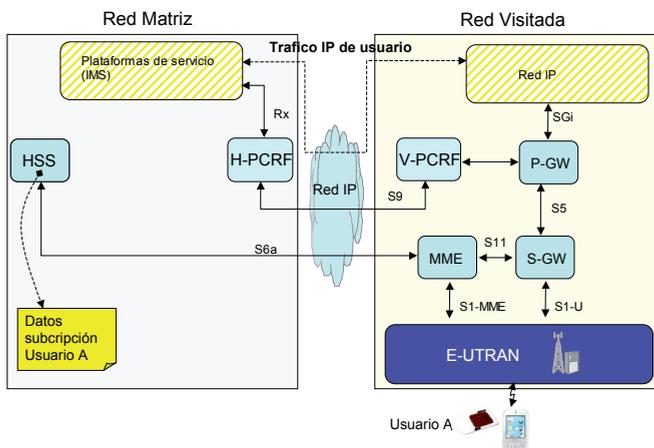


Figura 2.26 Encaminamiento de tráfico a través de la red visitada, con acceso a los servicios de la red matriz

2.7 IP Multimedia Subsystem (IMS)

El subsistema IMS proporciona los mecanismos de control necesarios para la provisión de servicios de comunicación multimedia basados en la utilización del protocolo IP a los usuarios de la red LTE. Así, el subsistema IMS se materializa mediante el despliegue de infraestructura constituida por una serie de elementos (servidores, bases de datos, pasarelas) que se comunican entre sí mediante diversos protocolos, fundamentalmente estándares del IETF, y que permiten gestionar la provisión de servicios tales como voz y video sobre IP, presencia y mensajería instantánea, servicios de llamadas en grupo, etc. El acceso de los terminales a dicha infraestructura se realiza a través de los servicios de conectividad IP que proporciona la red LTE. La provisión de servicios en redes de comunicaciones móviles a través de IMS pretende sustituir a medio-largo plazo los servicios equivalentes ofrecidos actualmente en modo circuito, perspectiva avalada por el hecho de que la nueva red de acceso E-UTRAN ya ha sido diseñada de forma que no proporciona acceso al dominio de circuitos, tal como se ha comentado en el apartado 2.4.

La adopción de protocolos desarrollados en IETF para la provisión de servicios multimedia en sistemas de comunicaciones móviles permite reducir considerablemente el ciclo de desarrollo de terminales y equipos, dada la relativa madurez tecnológica con que ya cuentan muchos de estos protocolos en otros contextos (e.g., Internet, redes de área local, etc.). Además, su adopción facilita en gran medida la interconexión de diferentes redes de telecomunicación ya sean móviles o fijas, públicas o privadas, que también se basan en dichos protocolos. Este enfoque ha ido cobrando fuerza en organizaciones como 3GPP, a diferencia del planteamiento adoptado en el desarrollo normativo de los primeros sistemas como GSM en ETSI, donde los protocolos utilizados se diseñaron específicamente para GSM. En particular, 3GPP escogió el protocolo SIP (Session Initiation Protocol) definido en la recomendación RFC3261 de IETF como protocolo de base para soportar la señalización asociada al subsistema IMS.

El subsistema IMS [38] tiene sus orígenes en el conjunto de especificaciones correspondientes a la *Release* 5 del sistema UMTS. Posteriormente, el ámbito de aplicación del IMS se ha extendido a otras tecnologías de red, tanto móviles (LTE, redes 3GPP2, Mobile WiMAX, etc.), como fijas (ADSL, cable, etc.). El modelo de provisión de servicios en base al subsistema IMS se estructura en tres capas: transporte, control y aplicación, tal como se muestra en la Figura 2.27. La capa de transporte representa la infraestructura de red IP, dependiente de la tecnología de acceso, que proporciona el encaminamiento de los flujos IP entre terminales y demás elementos de la red. Por tanto, el servicio de conectividad IP que ofrece la red LTE constituye claramente una realización válida de la capa de transporte en este modelo de provisión de servicios. En la capa de control se ubican los elementos especializados en la gestión de sesiones tales como los servidores de señalización SIP, así como otros elementos específicos para la interacción con redes telefónicas convencionales (pasarelas VoIP, controladores, etc.). Esta capa de control es la que se materializa mediante el subsistema IMS. Tal como se ilustra en la Figura 2.27, es importante destacar que los servicios de conectividad IP proporcionados por la capa de transporte pueden ser controlados desde la capa de control. Por último, en la capa de

aplicación residen los servidores de aplicación que albergan la lógica y datos asociados a los diferentes servicios proporcionados a través de IMS (e.g., servicio de mensajería instantánea y presencia, etc.). En la capa de aplicación también se pueden encontrar elementos ligados a otras plataformas de servicios como redes inteligentes y pasarelas *Parlay/OSA* [39]. A través de estas plataformas se posibilita la provisión de servicios desde proveedores de aplicaciones externos, denominados como *Application Service Providers, ASPs*.

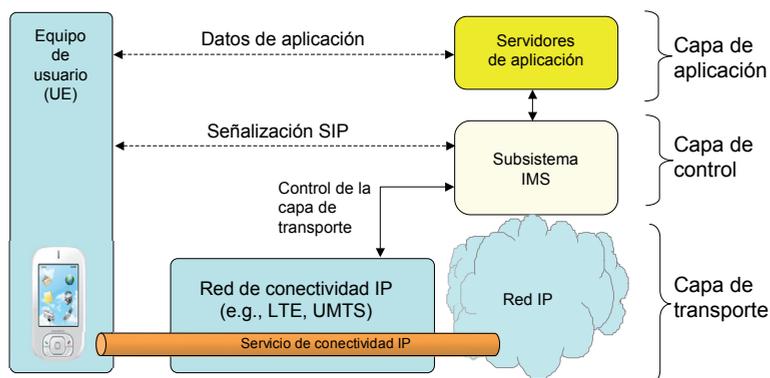


Figura 2.27 Modelo de provisión de servicios en base al subsistema IMS

El establecimiento y liberación de sesiones a través del IMS se basa en el protocolo de señalización SIP complementado con una serie de extensiones definidas por el 3GPP. Nótese que, tal como se ilustra en la Figura 2.27, un terminal conectado a través de, por ejemplo, LTE, utiliza SIP para su interacción con IMS. SIP es un protocolo que se concibió en un principio para controlar el establecimiento y liberación de sesiones multimedia (telefonía, videoconferencia, etc.) sobre redes IP entre dos o más participantes. Gracias a su flexibilidad, actualmente el ámbito de aplicación de SIP abarca una gama de aplicaciones mucho más extensa: mensajería instantánea y presencia, juegos distribuidos, control remoto de dispositivos, etc. Se trata pues de un protocolo en auge, ampliamente respaldado por organismos de normalización y por la industria, y en continua evolución. Además de SIP, en el IMS se emplean otros protocolos de señalización, tales como Diameter (visto en el apartado 2.6.3.3 en el contexto de la red troncal EPC), y MEGACO/H.248, para escenarios de comunicación en los que intervienen pasarelas de medios (por ejemplo, pasarelas VoIP para interconectar redes IP con redes telefónicas convencionales).

La especificación de la arquitectura funcional del subsistema IMS se aborda en el documento 3GPP 23.228 [38]. En aras a proporcionar al lector una visión completa de todos los componentes que integran un sistema LTE, a continuación se describen los principales elementos que forman parte del subsistema IMS, sin pretender alcanzar el mismo nivel de detalle seguido en la descripción de E-UTRAN y EPC. El lector interesado en una descripción más extensa del subsistema IMS puede encontrar una buena referencia en el libro [40].

En la Figura 2.28 se ilustra una arquitectura simplificada de los principales componentes

funcionales que integran el subsistema IMS. El núcleo del subsistema IMS lo forman las entidades denominadas CSCF (*Call Session Control Function*). Básicamente, se trata de servidores SIP que proporcionan las siguientes funciones:

- El *Serving CSCF* (S-CSCF) actúa como el nodo central de la señalización en sesiones IMS. El S-CSCF actúa como servidor de registro SIP (*SIP Registrar*) de forma que, cualquier mensaje de señalización SIP dirigido al usuario mediante la dirección SIP correspondiente, siempre termina pasando por el S-CSCF donde el usuario se encuentra registrado. S-CSCF puede proporcionar por si solo algunos servicios al usuario (e.g., servicio de redireccionamiento de llamadas, listas de marcado, etc.) o bien encaminar la señalización SIP hacia los servidores de aplicación correspondientes (e.g., servidores que implementan un servicio de mensajería instantánea).
- El *Proxy CSCF* (P-CSCF) es un servidor SIP que actúa como la puerta de entrada al subsistema IMS desde la red de conectividad IP: toda su señalización SIP de los terminales LTE/UMTS transcurre a través de este servidor. Entre otras funciones, el P-CSCF es el elemento que interacciona con las funciones de PCC (*Policy and Charging Control*) de la red de conectividad (a través de la entidad PCRF, tal como se ha descrito en el apartado 2.6). Por tanto, a través del P-CSCF, el subsistema IMS puede controlar la operativa de la capa de transporte (i.e., servicios portadores EPS en el caso del sistema LTE).
- El *Interrogating CSCF* (I-CSCF) es un servidor SIP que actúa como puerta de entrada de la señalización SIP proveniente de redes externas. En este sentido, la dirección IP del servidor I-CSCF es la que está incluida en los servidores de nombres de las redes externas (e.g., servidores DNS de Internet) de forma que la resolución de nombres de dominio (e.g., usuario@operadorLTE.com) conduce al envío de los mensajes SIP a la dirección IP del servidor I-CSCF quien, posteriormente, re-dirige los mensajes a los servidores S-CSCF correspondientes.

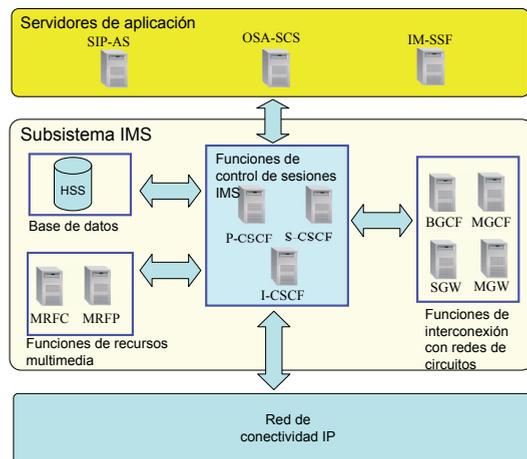


Figura 2.28 Arquitectura simplificada del subsistema IMS

Los servidores CSCF tienen acceso a la base de datos HSS. En dicha base de datos, tal como se ha descrito en el apartado 2.6.2.4, junto con la información necesaria para acceder a las redes 3GPP también se almacena información necesaria para soportar sesiones multimedia sobre IMS (e.g., servidor S-CSCF asignado a cada usuario).

El subsistema IMS contempla también mecanismos que permiten la interoperabilidad de los servicios IMS con servicios equivalentes ofrecidos a través de redes de conmutación de circuitos como la red telefónica convencional. De esta forma, un usuario de la red LTE puede establecer una llamada de voz a través del subsistema IMS con un usuario de la red telefónica conmutada. En una llamada de estas características, la transmisión de voz del usuario LTE se soporta mediante una solución de voz sobre IP a través del servicio de conectividad proporcionado por la red LTE, y el subsistema IMS se encargaría de realizar la conversión de códecs necesaria (e.g., terminal con un códec *Adaptive Media Rate*, AMR, y circuitos con G.711) y el establecimiento del circuito telefónico. Las funciones de interconexión con redes de circuitos se realizan a través de las denominadas pasarelas de medios (*Media Gateway*, MGW) que proporcionan la conectividad entre el plano de transporte de la red de paquetes (conexiones voz sobre IP) y los circuitos (e.g., canalizaciones de 64kbps). La pasarela de medios se controla desde un nodo de control (*Media Gateway Controller Function*, MGCF) a través del protocolo H.248 [40]. El controlador de la pasarela de medios es donde se realiza la conversión de la señalización SIP a la señalización utilizada en la red telefónica (e.g., *ISDN Signalling User Part*, ISUP).

El subsistema IMS integra también un conjunto de funciones de recursos multimedia. A modo de ejemplo, estas funciones permiten, bajo control de los servidores SIP, poner locuciones o tonos a usuarios de los servicios IMS. También permiten realizar la agregación de tráfico en sesiones SIP con más de dos interlocutores así como conversiones de códecs. Estas funciones se soportan en las entidades de red denominadas como *Media Resource Function Controllers* (MRFCs) y *Media Resource Function Processor* (MRFPs). El MRFC actúa como un agente de usuario SIP y controla los recursos del MRFP vía una interfaz H.248.

Finalmente, la lógica y datos de los servicios proporcionados mediante el subsistema IMS se ubica en los servidores de aplicación, tal como se ilustra en la Figura 2.28. Básicamente, un servidor de aplicación (denominado como *Application Server*, AS, en el contexto del subsistema IMS) es el servidor que contiene la lógica y ejecuta los servicios SIP. El sistema IMS identifica tres tipos diferentes de servidores de aplicación: SIP AS (*SIP Application Server*), OSA-SCS (*Open Service Access-Service Capability Server*) y IM-SSF (*IP Multimedia Service Switching Function*). El servidor de aplicaciones SIP (SIP AS) representa un servidor SIP genérico encargado de la provisión de un determinado servicio final. Es importante destacar que la especificación de servicios finales basados en SIP queda fuera del ámbito de 3GPP. En cualquier caso, con el objetivo de garantizar la interoperabilidad de diferentes servicios entre plataformas IMS, la organización *Open Mobile Alliance*⁶ (OMA) se ocupa de establecer el marco funcional apropiado que facilite el desarrollo de servicios IMS. El objetivo de OMA es especificar habilitadores de servicio ("*Service Enablers*") con la finalidad de fomentar la adopción global de servicios de

⁶ <http://www.openmobilealliance.org>

datos en redes móviles y garantizar la interoperabilidad de servicios de datos en redes móviles entre dispositivos, ubicaciones, proveedores de servicio, operadores móviles y redes. A modo de ejemplo, OMA especifica la arquitectura funcional de servicios tales como Mensajería Instantánea y Presencia, *Push to Talk Over Cellular* (PoC), y servicios *Push* sobre IMS [41]. Mediante el servidor de aplicaciones OSA-SCS, el subsistema IMS también ofrece la posibilidad de interactuar con el entorno de provisión de servicios *OSA/Parlay* [39]. Asimismo, a través del servidor IM-SSF el subsistema IMS también permite que los servicios IMS puedan interactuar con los mecanismos de red inteligente que soportan las redes 3GPP (servicios CAMEL, *Customised Applications for Mobile network Enhanced Logic*, [42]).

2.8 Equipos de usuario

El equipo de usuario es el equipo que permite a los usuarios del sistema LTE acceder a los servicios de la red LTE a través de la interfaz radio. La arquitectura funcional de un equipo de usuario en el sistema LTE es la misma que en su momento se definió para los sistemas GSM y que se adaptó posteriormente para UMTS. La arquitectura funcional de un equipo de usuario GSM/UMTS/LTE se muestra en la Figura 2.29. El equipo de usuario (*User Equipment, UE*) contiene dos elementos básicos: un módulo de suscripción del usuario (SIM/USIM) y el equipo móvil propiamente dicho (*Mobile Equipment, ME*). Adicionalmente, las funciones del equipo móvil se agrupan en dos entidades funcionales: la terminación móvil (*Mobile Terminal, MT*) y el equipo terminal (*Terminal Equipment, TE*). A continuación se describen cada uno de ellos.

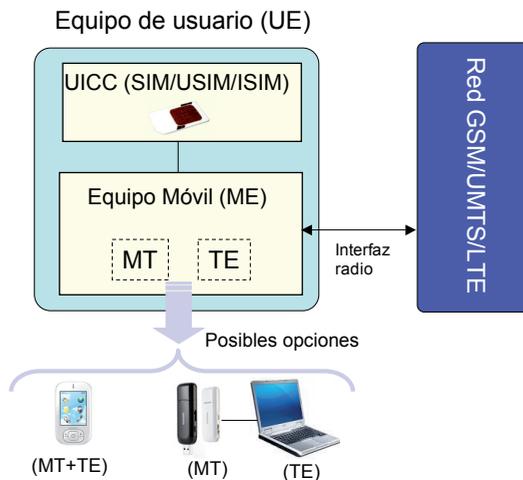


Figura 2.29 Equipo de usuario

El módulo de suscripción de usuario se materializa mediante el uso de una tarjeta inteligente (UICC) que contiene la aplicación que gestiona los datos de suscripción de los usuarios

a las redes 3GPP. La aplicación utilizada para acceder a redes GSM se denomina SIM (Subscriber Identity Module), mientras que para el acceso a redes UMTS se conoce como UMTS SIM (USIM). Además de las aplicaciones SIM y USIM, la tarjeta UICC puede contener también una aplicación denominada ISIM (IP Multimedia Services Identity Module) para almacenar los datos necesarios para la operación de los servicios IMS (e.g., direcciones SIP del usuario). La SIM/USIM está asociada a un usuario y por tanto es quien le identifica en el sistema independientemente del equipo móvil utilizado. La separación entre USIM y ME facilita que un usuario pueda cambiar de terminal manteniendo su identidad. Muchas veces los términos tarjeta inteligente, SIM y USIM se utilizan de forma equivalente. Estrictamente, la tarjeta inteligente únicamente se refiere al circuito integrado y sus capacidades de comunicación y procesamiento (elementos hardware/software). Como entorno de procesamiento que es, una tarjeta inteligente puede ejecutar diferentes aplicaciones, y dos de ellas son la SIM y USIM. Dichas aplicaciones contienen una estructura de datos (e.g., IMSI del usuario, áreas de seguimiento, etc.) y un conjunto de rutinas específicas (e.g., soporte del mecanismo de autenticación). El diseño de estas aplicaciones se realiza de forma que se garantice compatibilidad con los sistemas previos. Así, una tarjeta con una aplicación USIM, o tarjeta USIM, puede ser utilizada para acceder a una red GSM. En el caso del sistema LTE, se ha mantenido la misma compatibilidad, y además, la aplicación correspondiente sigue denominándose USIM. No obstante, la USIM utilizada para acceder a un sistema LTE, extiende la aplicación USIM de UMTS mediante la inclusión de los parámetros adicionales necesarios para gestionar el acceso a través de LTE (e.g., las nuevas claves de seguridad soportadas en LTE). La descripción de la USIM con soporte para LTE se especifica en TS 31.102 [43].

El equipo móvil (ME) integra las funciones propias de comunicación con la red celular así como las funciones adicionales que permiten la interacción del usuario con los servicios de la red (nótese que un usuario puede ser una persona o bien otro dispositivo electrónico). En este sentido, de cara a introducir un cierto grado de flexibilidad en el diseño del equipo móvil, se ha definido una interfaz que permite que exista una separación física entre el equipo que alberga las funciones propias de la comunicación (MT) y el equipo que se ocupa de la interacción con el usuario (TE), y que puede contener multitud de aplicaciones adicionales no directamente relacionadas con el sistema de comunicaciones móviles. Dicha interfaz consiste en la especificación del conjunto de comandos AT que permiten acceder a los servicios de la red (e.g., establecimiento de una conexión en la red LTE) soportados en el MT desde el TE [44]. En cualquier caso, la separación entre TE y MT es opcional y simplemente pretende flexibilizar el desarrollo de equipos de usuario más complejos. Tal como se ilustra en la Figura 2.29, un ejemplo claro de equipo de usuario que explota la interfaz basada en comandos AT es la utilización de un portátil (TE) y un módem GSM/UMTS/LTE (MT).

La complejidad y prestaciones que puede presentar un equipo de usuario LTE se clasifica en un conjunto de categorías definidas por 3GPP. Las categorías, denominadas formalmente como *ue-Category* en las especificaciones del 3GPP, se establecen en función de la capacidad de transmisión del terminal. Así, las categorías tienen en cuenta el grado de soporte de los terminales LTE de mecanismos de multiplexación espacial con múltiples antenas, el

uso de determinadas modulaciones, el tamaño en bytes de las colas de transmisión, etc. En particular, en la primera versión del sistema LTE (*Release 8*) se han definido 5 categorías de terminal cuyas características se describen en el Capítulo 5. Además de la categoría del terminal en términos de capacidad de transmisión, un equipo de usuario también admite diferentes implementaciones en base, por ejemplo, a las bandas de frecuencias soportadas y a la integración en el propio terminal de otras tecnologías de acceso radio (e.g., terminales multimodo 3GPP/802.11). La especificación detallada de las diferentes capacidades que pueden contemplarse en los terminales LTE se proporciona en 3GPP TS 36.306 [45].

2.9 Referencias bibliográficas

- [1] 3GPP 23.002, "Network architecture".
- [2] Especificaciones NGN de TISPAN disponibles on-line en <http://www.etsi.org/tispan/>.
- [3] Next Generation Networks Global Standards Initiative (NGN-GSI). Website: <http://www.itu.int/ITU-T/ngn/introduction.html>.
- [4] 3GPP 36.300, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2".
- [5] 3GPP 36.401, "Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Architecture description".
- [6] 3GPP TS 36.410, "S1 General Aspects and Principles".
- [7] 3GPP TS 36.420, "X2 General Aspects and Principles".
- [8] 3GPP TS 36.413, "S1 Protocol Specification".
- [9] IETF RFC4960, "Stream Control Transmission Protocol".
- [10] 3GPP TS 36.423, "X2 Protocol Specification".
- [11] 3GPP TS 25.401, "UTRAN overall description".
- [12] Bosch, H.G.P. et al., "Flat Cellular (UMTS) Networks", Wireless Communications and Networking Conference (WCNC), marzo 2007, Kowloon, China.
- [13] Rajeev Agrawal and Anand Bedekar, "Network Architectures for 4G: Cost Considerations", IEEE Communications Magazine, diciembre 2007.
- [14] 3GPP TS 23.401, "General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access".
- [15] 3GPP TS 23.402, "Architecture enhancements for non-3GPP accesses".
- [16] 3GPP TS 23.008, "Organization of subscriber data".
- [17] 3GPP TS 29.061, "Interworking between the Public Land Mobile Network (PLMN) supporting packet based services and Packet Data Networks (PDN)".

- [18] "TCP/IP Illustrated, Volume 1: The Protocols", W. Richard Stevens, Editorial Addison-Wesley Professional, 1994.
- [19] "Mobile Inter-networking with IPv6: Concepts, Principles and Practices", Rajeev S. Koodli y Charles E. Perkins, Editorial Wiley-Interscience, 2007.
- [20] "UMTS Networks: Architecture, Mobility and Services", H. Kaaranen et al., Editorial John Wiley & Son, segunda edición 2005.
- [21] 3GPP TS 29.274, "3GPP Evolved Packet System (EPS); Evolved General Packet Radio Service (GPRS) Tunnelling Protocol for Control plane (GTPv2-C); Stage 3".
- [22] 3GPP TS 29.060, "GPRS Tunnelling Protocol (GTP) across the Gn and Gp interface".
- [23] IETF 2865 (junio 2000), "Remote Authentication Dial In User Service (RADIUS)".
- [24] 3GPP TS 23.275, "Proxy Mobile IPv6 (PMIPv6) based Mobility and Tunnelling protocols; Stage 3".
- [25] 3GPP TS 29.272, "Evolved Packet System (EPS); Mobility Management Entity (MME) and Serving GPRS Support Node (SGSN) related interfaces based on Diameter protocol".
- [26] 3GPP TS 29.281, "General Packet Radio System (GPRS) Tunnelling Protocol User Plane (GTPv1-U)".
- [27] 3GPP TS 23.203, "Policy and charging control architecture".
- [28] 3GPP TS 32.240, "Telecommunication management; Charging management; Charging architecture and principles".
- [29] 3GPP TS 29.214, "Policy and charging control over Rx reference point".
- [30] 3GPP TS 29.215 "Policy and Charging Control (PCC) over S9 reference point".
- [31] 3GPP TS 29.212, "Policy and charging control over Gx reference point".
- [32] IETF RFC 3588, "Diameter Base Protocol".
- [33] IETF RFC 4005, "Diameter Network Access Server Application".
- [34] IETF RFC 4006, "Diameter Credit-Control Application".
- [35] IETF RFC 5213, "Proxy Mobile IPv6".
- [36] IETF RFC 2784, "Generic Routing Encapsulation (GRE)".
- [37] 3GPP TS 24.301, "Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); Stage 3".
- [38] 3GPP TS 23.228, "IP Multimedia Subsystem (IMS); Stage 2".
- [39] "Parlay/OSA: From Standards to Reality", Musa Unmehopa, Kumar Vemuri, Andy Bennett, Editorial John Wiley & Sons, 2006.

- [40] The 3G IP Multimedia Subsystem (IMS): Merging the Internet and the Cellular Worlds, Gonzalo Camarillo y Miguel Ángel García, Editorial Wiley, Tercera Edición 2008.
- [41] Open Mobile Alliance, http://www.openmobilealliance.org/technical/released_enablers.aspx.
- [42] 3GPP TS 29.278, "Customized Applications for Mobile network Enhanced Logic (CAMEL); CAMEL Application Part (CAP) specification for IP Multimedia Subsystems (IMS)".
- [43] 3GPP TS 31.102, "Characteristics of the Universal Subscriber Identity Module (USIM) application".
- [44] 3GPP TS 27.007, "AT command set for User Equipment (UE)".
- [45] 3GPP TS 36.306, "Evolved Universal Terrestrial Radio Access (E-UTRA); "UE Radio Access Capabilities".



III

**Marco de gestión de sesiones,
movilidad y seguridad en LTE**

3 Marco de gestión de sesiones, movilidad y seguridad en LTE

3.1 Introducción

Este capítulo proporciona una descripción de los conceptos fundamentales de un sistema LTE en cuanto a la gestión de sesiones, movilidad y seguridad.

La gestión de sesiones se refiere a la gestión del servicio de conectividad IP que ofrece una red LTE. Este servicio de conectividad IP se articula en base a los conceptos de *conexión PDN* y *servicio portador EPS* que se describen en detalle en este capítulo. Tal como se verá, el servicio de conectividad IP de una red LTE es capaz de ofrecer diferentes niveles de calidad de servicio (QoS) y puede ser gestionado mediante mecanismos de control de las políticas de uso de la red (i.e., *policy control*) que permiten la interacción del servicio de conectividad proporcionado por LTE con las plataformas que sustentan los servicios finales (e.g., IMS). Los conceptos, funcionalidades y procedimientos más relevantes asociados a la gestión de sesiones en LTE se abordan en el apartado 3.2.

La gestión de movilidad se refiere a los mecanismos con los que cuenta el sistema LTE para que los usuarios puedan acceder y recibir servicios desde cualquier ubicación geográfica donde el sistema disponga de cobertura. Asimismo, la gestión de movilidad también abarca los mecanismos utilizados en el sistema LTE para poder mantener las conexiones de sus usuarios activas aun cuando éstos puedan cambiar de estación base debido a su movilidad (i.e., mecanismos de *handover*). La gestión de movilidad en LTE se analiza en el apartado 3.3.

La gestión de seguridad concierne a cómo la red LTE autentica y autoriza el uso de sus servicios a los usuarios así como cuáles son los mecanismos utilizados para proporcionar confidencialidad e integridad de la información enviada tanto en la interfaz radio como en otras interfaces entre equipos de red. Los componentes y procedimientos básicos de la gestión de seguridad en LTE se cubren en el apartado 3.4.

3.2 Gestión de sesiones

3.2.1 Servicio de conectividad IP: conexiones PDN

El sistema LTE proporciona a los usuarios un servicio de conectividad IP a una o múltiples redes de paquetes externas. El servicio de conectividad IP permite que un terminal LTE pueda intercambiar información (i.e., paquetes IP) con otros equipos de la red IP externa remota como si el terminal LTE se encontrara conectado físicamente a dicha red (i.e., el terminal es “visible” en la red remota externa mediante una dirección IP propia que le ha sido asignada a través del sistema LTE). El servicio de conectividad IP proporcionado por LTE entre el equipo de usuario y una red externa IP se denomina “conexión PDN”.

Las redes de paquetes externas a las que puede accederse a través de la red LTE pueden ser redes públicas como Internet o bien redes privadas como una intranet corporativa, la red de un proveedor de acceso a Internet (ISP) o bien una red interna del propio operador de la

red LTE para la provisión, por ejemplo, de servicios IMS. En las especificaciones del 3GPP, las posibles redes externas de paquetes accesibles desde LTE se denominan, sin distinción alguna, como *Packet Data Networks* (PDNs), y por este motivo se habla de conexión PDN en lugar de simplemente conexión IP. La denominación de PDN es herencia de GPRS donde se acuñó este término. En GPRS, el servicio de conectividad ofrecido contempla la transferencia de otros tipos de paquetes además de IP como, por ejemplo, paquetes PPP (*Point-to-Point Protocol*)¹ o paquetes X.25. Por ello, en GPRS se habla de forma genérica de conexiones PDN, en lugar de conexiones IP, si bien es cierto que el servicio de conectividad IP es el caso predominante en los despliegues actuales de este tipo de redes. A diferencia de GPRS, el sistema LTE únicamente proporciona conexiones a redes IP (LTE no soporta la transmisión de otro tipos de paquetes). No obstante, en las especificaciones del sistema LTE sigue utilizándose el término “conexión PDN” utilizado en GPRS para referirse al servicio de conectividad IP ofrecido por la red.

Una conexión PDN en el sistema LTE se caracteriza siempre por una dirección IP única a través de la cual el equipo de usuario opera en la red externa. La dirección IP de una conexión PDN puede ser IPv4, IPv6 o ambas (una conexión puede estar asociada simultáneamente a una dirección IPv4 y a una IPv6). En la Figura 3.1 se ilustra el servicio de conectividad IP (i.e., conexión PDN) proporcionado por la red LTE.

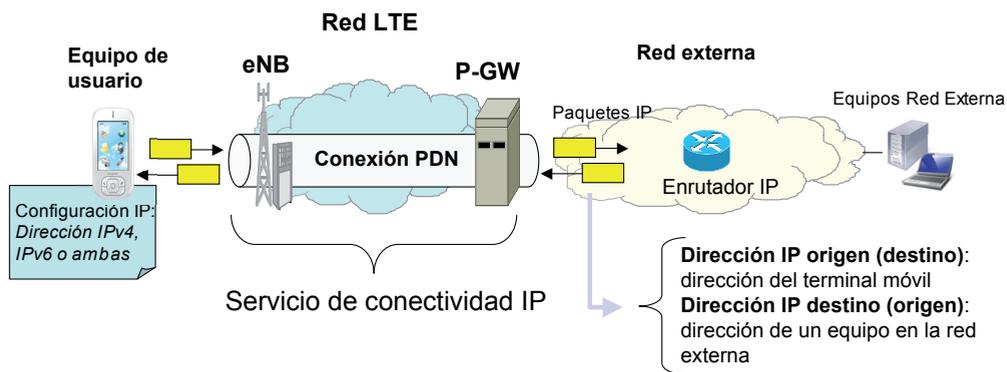


Figura 3.1 Servicio de conectividad IP de la red LTE: conexión PDN

Las redes externas a las que una red LTE puede proporcionar acceso se identifican mediante una etiqueta denominada *Access Point Name* (APN). El APN se compone de un identificador del operador de la red LTE (un código de operador y de país) y un identificador específico de la red externa a la que se proporciona acceso (e.g., “servicios-ims”, “internet”, “red-corporativa-1”, etc.). De esta forma, cuando se establece una conexión PDN entre un equipo de usuario y una red externa, la red LTE utiliza el parámetro APN para determinar la pasarela P-GW o pasarelas P-GW de la red LTE que van a participar en la provisión de dicha conexión PDN.

¹ El protocolo PPP es en realidad un protocolo de nivel de enlace que permite transportar de forma encapsulada diferentes protocolos de nivel de red (e.g., IP sobre PPP) y que proporciona funciones adicionales para la autenticación y la asignación de direcciones de red. El protocolo PPP es ampliamente utilizado para la conexión a Internet con un módem telefónico.

Un equipo de usuario puede establecer múltiples conexiones PDN simultáneas, a través de la misma o varias pasarelas P-GW de la red LTE. El conjunto de redes externas a las que tiene acceso un determinado usuario LTE se controla a través de su subscripción, donde se indica el conjunto de identificadores APNs autorizados. En LTE, a partir del momento en que un usuario se registra en la red LTE, se establece, como mínimo, una conexión PDN. La red externa a la que proporciona acceso la conexión PDN inicial la puede decidir el propio usuario mediante el envío del correspondiente APN o bien puede utilizarse un valor de APN por defecto que la red LTE guarde en el perfil de subscripción de dicho usuario. Las conexiones PDN adicionales, si cabe, las inicia el equipo de usuario una vez ya está registrado y dispone de una conexión PDN inicial (el registro y el establecimiento de la primera conexión se realiza de forma combinada mediante el procedimiento denominado *Network Attach* que se analizará en detalle en el apartado 3.2.5.1). La desconexión de una conexión PDN puede iniciarla tanto la propia red LTE (debido, por ejemplo, a un cambio en la subscripción o a una falta de recursos) como el propio usuario.

A modo ilustrativo, en la Figura 3.2 se representa un equipo de usuario que tiene establecidas tres conexiones PDN con sendas redes externas. Tal como se ha comentado en párrafos anteriores, cada conexión PDN se caracteriza por un identificador APN (que indica a qué red externa se proporciona acceso, y por extensión, desde qué pasarela P-GW) y la dirección IP utilizada por el terminal LTE en dicha red externa. Nótese, por tanto, que un terminal LTE puede tener asignadas múltiples direcciones IP, una para cada conexión PDN establecida.

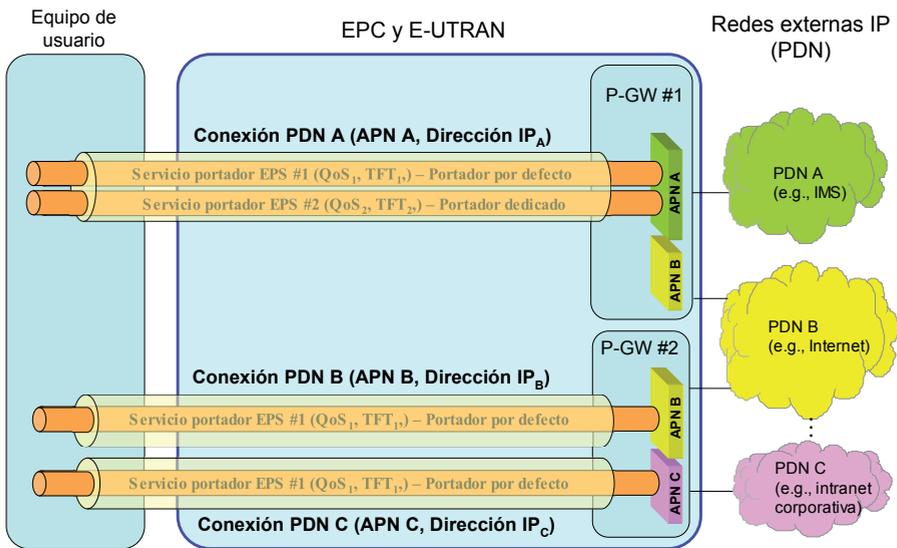


Figura 3.2 Conexiones PDN y servicios portadores EPS

Según la Figura 3.2, mediante la conexión PDN A, el usuario tiene acceso a servicios proporcionados en una plataforma IMS. Por otro lado, la conexión PDN B proporciona al usuario acceso a Internet. En este ejemplo, el acceso a Internet como red externa se indica median-

te el identificador APN B y puede realizarse tanto a través de la pasarela P-GW#1 como de P-GW#2, tal como se ilustra en la Figura 3.2. Cuando existen múltiples pasarelas P-GW que proporcionan acceso a una misma red externa, la elección de la pasarela más apropiada la determinaría la red LTE en base a diferentes aspectos tales como balanceo de cargas, espacio de direcciones utilizado en cada pasarela, diferenciación del servicio de acceso a Internet, etc. Finalmente, según la Figura 3.2, el equipo de usuario mantiene una tercera conexión PDN a una red privada que se identifica mediante el APN C y cuyo acceso se logra a través de la pasarela P-GW#2. En caso de que hubiera múltiples conexiones PDN asociadas al mismo APN, estas deberían ser proporcionadas por la misma pasarela. El soporte de múltiples conexiones PDN desde el mismo equipo de usuario es opcional en el sistema LTE. En la Figura 3.2, también puede verse como cada conexión PDN está compuesta por un conjunto de servicios portadores EPS (*EPS bearer services*). El concepto de servicio portador EPS, su estructura interna y su relación con las conexiones PDN, se describen a continuación.

3.2.2 Servicio portador EPS

El servicio de conectividad IP proporcionado por el sistema LTE es un servicio que soporta calidad de servicio (*Quality of Service*, QoS). Así, el trato que reciben los paquetes IP de una determinada conexión PDN en términos de, por ejemplo, tasa de transferencia en bits/s, retardo de transmisión y tasa de pérdidas de paquetes, puede adaptarse a las necesidades de transmisión de los servicios finales a los que accede el usuario. En este contexto, es importante tener en cuenta que a través del sistema LTE pueden proporcionarse servicios finales de muy diferente índole que no requieren las mismas prestaciones del servicio de transmisión (e.g., transmisión de audio y video en tiempo real, servicios de mensajería, etc.). Por tanto, la adaptación de las prestaciones de QoS de las conexiones PDN a las características de los servicios finales permite al sistema LTE proporcionar una buena experiencia de uso a los usuarios a la vez que posibilita una gestión eficiente de los recursos (i.e., no se reservan más recursos de transmisión de los estrictamente necesarios para satisfacer los objetivos de calidad de servicio). Nótese que la configuración del comportamiento de QoS, aparte de ser dependiente del servicio final al que accede un usuario, también puede permitir al operador de la red LTE llevar a cabo unas determinadas estrategias de negocio en base a la diferenciación de usuarios (e.g., mediante una suscripción "gold" un usuario puede recibir una tasa de transferencia superior en el acceso a Internet).

La forma de gestionar la calidad de servicio en el sistema LTE se estructura en torno a la definición de lo que se denomina servicio portador *EPS* (*EPS Bearer Service*). Un servicio portador EPS es un servicio de transferencia de paquetes IP que tiene asociados unos parámetros de QoS y la plantilla o filtro de paquetes (denominado como *Traffic Flow Template*, TFT) utilizada para seleccionar el flujo de paquetes IP que debe recibir dicho trato de QoS. En este sentido, el servicio portador EPS constituye la unidad mínima de resolución para la provisión de QoS: todos los paquetes IP que fluyen en la red asociados a un mismo servicio portador EPS reciben el mismo trato de QoS.

La transmisión de paquetes IP en las conexiones PDN se articula en base al establecimiento de, como mínimo, un servicio portador EPS, tal como se ilustra en la Figura 3.2. Así, en cada conexión PDN siempre existe un servicio portador EPS por defecto activo por el que se enviaría todo el tráfico IP de usuario sin distinción alguna (e.g., señalización SIP y datos de las diferentes aplicaciones en curso). Opcionalmente, en aras a poder proporcionar un trato de QoS específico a un determinado flujo de paquetes IP (e.g., paquetes IP correspondientes a una aplicación de videoconferencia) pueden activarse servicios portadores EPS adicionales al portador por defecto, denominados como servicios portadores EPS dedicados. De esta forma, si en una conexión PDN existen servicios portadores EPS dedicados activos, el tráfico seleccionado (mediante el TFT) se envía a través de ellos y el resto se cursa a través del portador EPS por defecto.

Los parámetros de QoS del servicio portador por defecto vienen fijados por la subscripción del usuario y, como tal, se encuentran en la información almacenada en la base de datos HSS del sistema LTE. En caso de que el usuario acceda a servicios IMS a través del servicio portador por defecto de una conexión PDN, los valores de QoS utilizados deben ser adecuados para la transferencia de señalización SIP entre el equipo de usuario y los servidores de control de la plataforma IMS. El servicio portador por defecto permanece activado durante la vigencia de la conexión PDN y su desactivación conduce a la terminación de la conexión PDN. En cambio, si se hace uso de servicios portadores EPS dedicados, éstos pueden activarse/modificarse/desactivarse al inicio o bien en el transcurso de una conexión PDN. La existencia de un servicio portador EPS dedicado suele estar vinculada a la existencia de un servicio final que requiere un trato específico de QoS (e.g., el establecimiento de un servicio videoconferencia puede conllevar la activación de un servicio portador EPS dedicado con valores de QoS adecuados para la transmisión de los paquetes de voz durante el tiempo que dure el servicio). Es importante destacar que todos los servicios portadores EPS de una misma conexión PDN comparten los mismos parámetros de conectividad IP (i.e., misma dirección IP). Tal como puede observarse en el ejemplo de Figura 3.2, en las tres conexiones PDN existe un servicio portador EPS por defecto y, en este ejemplo, únicamente en la conexión PDN A, hay activado un servicio portador EPS dedicado. Por tanto, en las conexiones PDN B y C, todo el tráfico recibe el mismo trato de QoS (el del servicio portador por defecto) mientras que en la conexión PDN A se aplican dos tratos de QoS diferentes (e.g., uno al servicio de videoconferencia y otro al resto de paquetes). A continuación se analiza más en detalle la implementación del servicio portador EPS dentro de la red LTE. Los parámetros de QoS del servicio portador EPS se abordan posteriormente en el apartado 3.2.3.

3.2.2.1 Estructura de un servicio portador EPS

Tal como se observa en la Figura 3.2, el servicio portador EPS se extiende desde el terminal hasta la pasarela P-GW de la red LTE proporcionando QoS *end-to-edge*². Los componentes

² QoS end-to-edge hace referencia a la provisión de un comportamiento de QoS entre el equipo de usuario desde donde se accede a los servicios ("end") y el punto de interconexión de LTE con la red externa ("edge"). Nótese la diferencia con un modelo QoS end-to-end, donde el comportamiento de QoS debe estar determinado entre los dos extremos de un servicio (e.g., entre el equipo de usuario y un servidor de videostreaming ubicado en una red externa a LTE). El sistema LTE no especifica un modelo de QoS end-to-end que cubra la transmisión en la red externa.

en que se estructura un servicio portador se ilustran más detalladamente en la Figura 3.3, donde se representa un escenario con dos servicios portadores EPS establecidos (A y B). A continuación se describen los diferentes componentes.

La implementación de un servicio portador EPS requiere que, en cada nodo de la red LTE donde se procesa el plano de usuario (i.e., equipo de usuario, eNB, S-GW y P-GW), se establezca un contexto³ de información con los parámetros de QoS pertinentes así como la información necesaria (filtros, etiquetas) que permita la identificación del flujo de datos asociado al servicio portador en las diferentes interfaces de la red. Así, cada servicio portador EPS tiene asociado un filtro de paquetes IP que se utiliza para seleccionar los paquetes que deben recibir el trato de QoS especificado. El filtro de paquetes recibe el nombre de TFT (*Traffic Flow Template*) y contiene atributos tales como: puertos de los protocolos de transporte, direcciones IP y máscaras de subred de los nodos remotos; puertos locales de los protocolos de transporte; y campos específicos de calidad de servicio en las cabeceras del protocolo IP (*Type of Service*, TOS, en IPv4 y *Traffic Class* y *Flow Label* en IPv6). El filtro de paquetes se instala en el equipo de usuario para filtrar el tráfico ascendente y en las pasarelas de red (P-GW o S-GW) para filtrar el tráfico dirigido al terminal. En la Figura 3.3 se ilustra la ubicación de los filtros de paquetes en el caso de considerar una interfaz basada en GTP entre S-GW y P-GW, donde el TFT en sentido descendente se aplica en la pasarela P-GW. En caso de que la interfaz entre S-GW y P-GW estuviera basada en PMIPv6 (véase el apartado 2.6.3.4 del Capítulo 2), los filtros TFT para el enlace descendente se ubicarían en la pasarela S-GW dado que el protocolo PMIPv6 no soporta la identificación de diferentes flujos de tráfico entre las pasarelas S-GW y P-GW (tal como se comenta en el apartado 2.6.3.4 del Capítulo 2, los túneles PMIPv6 entre pasarelas se establecen por conexión PDN, no por servicio portador EPS).

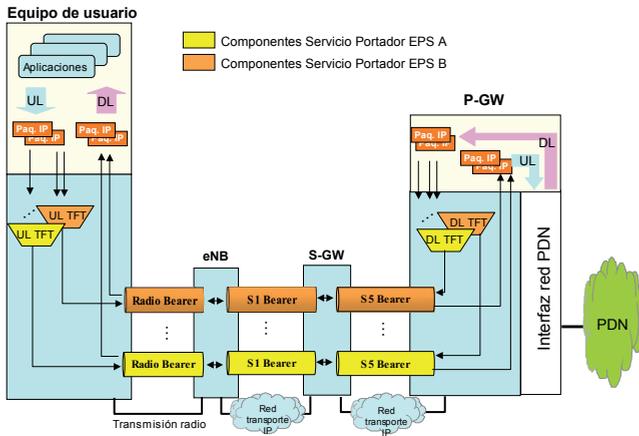


Figura 3.3 Componentes de los servicios portadores EPS

³ El contexto es la estructura de datos que utiliza una entidad de la red para almacenar información asociada a la operativa de un usuario. Dependiendo de la entidad de red en cuestión, el contexto contiene información de suscripción del usuario, capacidades del terminal e información relativa a la gestión de movilidad y de sesiones (e.g., configuración de los servicios portadores establecidos).

Una vez identificados los paquetes IP asociados a un servicio portador mediante los TFTs, éstos se transportan a través de la red de forma que resulta posible identificar el servicio portador EPS al que pertenece cada uno de los paquetes IP enviados en todos los equipos de red que sustentan el plano de usuario. Para ello, en la interfaz radio, los paquetes asociados a un servicio portador EPS se transmiten mediante un servicio portador radio (*Radio Bearer*, RB). Un servicio portador EPS se mapea unívocamente a un RB en la interfaz radio. En la interfaz S1 entre eNB y S-GW, el servicio portador EPS se mapea en un servicio portador S1 que básicamente queda especificado mediante la asignación de un identificador de túnel (TEID) del protocolo GTP en ambos extremos de la interfaz (véase apartado 2.6.3.1 del Capítulo 2). De forma similar, en la interfaz S5/S8 basada en GTP el mapeo del servicio portador EPS también consiste en su asociación con los identificadores de túnel que definen el servicio portador S5/S8. Tal como se ha comentado anteriormente, si la interfaz S5/S8 se basa en PMIPv6, la distinción entre los paquetes IP que pertenecen a diferentes servicios portadores EPS no se lleva a cabo en esta interfaz. En este caso, el servicio portador EPS queda implementado únicamente mediante el servicio portador radio y el servicio portador S1.

Los parámetros de QoS se establecen por servicio portador EPS y se mapean a los parámetros de QoS que soportan los servicios portadores integrantes (RB, S1 y S5/S8). En el caso de la interfaz radio, el cumplimiento de los parámetros de QoS se fundamenta en la configuración apropiada de la capa radio y de enlace (e.g., configuración de los mecanismos de retransmisiones RLC y H-ARQ) y en la utilización de mecanismos de *scheduling*, gestión activa de colas y control de la tasa⁴ de transferencia (bits/s) del servicio portador radio, tal como se analiza más en detalle en el Capítulo 6 en el contexto de funciones de gestión de recursos radio. En el caso de las interfaces S1 y S5/S8, los parámetros de QoS del servicio portador EPS deben relacionarse con los parámetros de QoS que empleen las redes de transporte IP que interconectan los diferentes equipos de la infraestructura de la red LTE (e.g., parámetros de un modelo de servicios Diffserv). En el siguiente apartado se describen en detalle los parámetros del modelo de QoS en el sistema LTE.

3.2.3 Modelo de QoS

El modelo de QoS “delimita” el comportamiento esperable del servicio de conectividad proporcionado por una red de transmisión de paquetes. La concreción del comportamiento esperable se materializa mediante la especificación de un conjunto de parámetros de QoS asociados a la obtención de unas determinadas prestaciones en términos de, por ejemplo, tasa (bits/s), retardos de transferencia y tasa de pérdida de paquetes. La forma de “delimitar” el comportamiento admite diferentes grados de determinismo que condicionan el tipo de mecanismos necesarios para poder aplicar el modelo de QoS. A modo de ejemplo, en redes IP basadas en una arquitectura de QoS Diffserv, el modelo de QoS se plantea en base a conseguir diferenciar tráfico (e.g., priorizar el envío de unos paquetes frente a otros) y no se contempla el poder establecer una cota de retardo máximo. Esto hace que la implementación del mo-

⁴ Mecanismo de *rate control*

delo Diffserv sea relativamente sencilla y escalable. En cambio, en una arquitectura de QoS IntServ, el modelo de QoS permite especificar una cota máxima del retardo de transferencia entre dos nodos de una red IP. La materialización de dicha cota conlleva la necesidad de reservar recursos de transmisión en cada uno de los nodos de la red, por lo que la aplicación de dicho modelo tiene problemas de escalabilidad. Un modelo de QoS también determina el grado de flexibilidad que ofrece un sistema para gestionar la capacidad de transmisión disponible. Esta claro que, en redes como LTE, planteamientos de sobredimensionado de la capacidad de transmisión de la red no son viables económicamente (el espectro es limitado y el coste de la infraestructura de red muy elevado). Por tanto, la porción de capacidad de transmisión asignada a un determinado servicio portador debe establecerse en base a las prestaciones mínimas exigibles para garantizar una buena experiencia de uso del servicio final.

El comportamiento esperable en términos de prestaciones de QoS de un servicio portador EPS depende del tipo de servicio final que se curse a través de dicho servicio portador. A modo de ejemplo, el comportamiento de QoS exigible será completamente diferente en caso de cursar un servicio de VoIP o un servicio de navegación HTTP. Asimismo, la cantidad de recursos de transmisión y la operativa de asignación dinámica de los mismos también será completamente diferente en ambos casos. Adicionalmente, los parámetros de QoS exigibles al servicio portador pueden también fijarse en función del tipo de usuario. De esta forma, el soporte de QoS puede explotarse también para diferenciar entre usuarios en función, por ejemplo, de las condiciones suscritas en la contratación del servicio. Bajo esta perspectiva, para un mismo servicio, el operador de la red LTE puede ofrecer un comportamiento del sistema diferente para, por ejemplo, usuarios de negocios frente a usuarios convencionales, usuarios de contrato frente a usuarios pre-pago, usuarios privilegiados en situaciones de emergencia (policía, bomberos), etc.

En LTE, el modelo de QoS utilizado para definir el comportamiento de un servicio portador EPS se basa en la especificación de un máximo de cuatro parámetros. Además de estos cuatro parámetros, el modelo de QoS se complementa con dos parámetros adicionales asociados a la suscripción de un usuario. En la Figura 3.4 se esquematiza el conjunto completo de parámetros de QoS considerado en el sistema LTE.

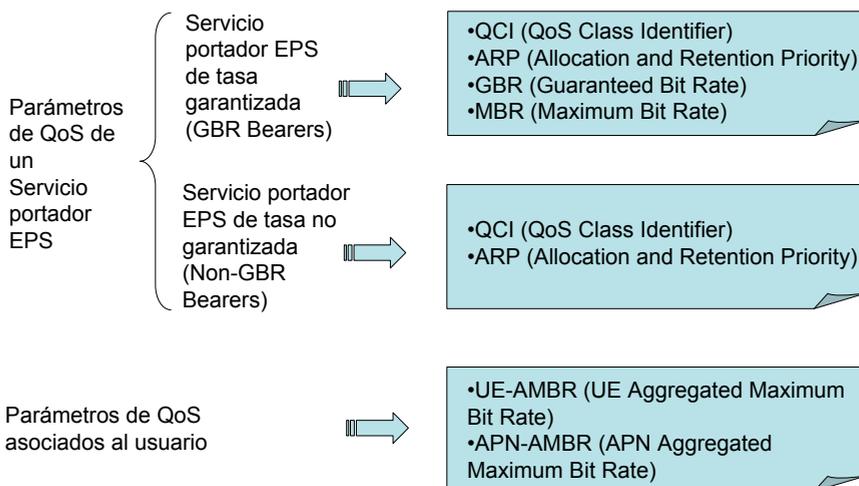


Figura 3.4 Parámetros de QoS en el sistema LTE

Cada servicio portador EPS siempre tiene asociados como mínimo dos parámetros: QCI (*QoS Class Identifier*) y ARP (*Allocation and Retention Priority*). De forma general, el parámetro QCI determina el comportamiento del plano de usuario del servicio portador EPS mientras que el parámetro ARP aplica a la operativa del plano de control. Adicionalmente, algunos servicios portadores denominados como servicios de tasa garantizada (*GBR Bearers*) especifican también un parámetro de tasa media garantizada (GBR) y otro de tasa máxima permitida (MBR). A continuación se describe más detalladamente el significado de dichos parámetros.

El QCI es un parámetro que representa una determinada clase de servicio o comportamiento de la red. El valor del QCI no indica de forma directa ninguna magnitud relacionada con las prestaciones de la red sino que simplemente se concibe como un puntero a una determinada clase de servicio⁵. De esta forma, la selección de un valor de QCI para un servicio portador EPS implica la utilización de una serie de parámetros específicos en cada uno de los nodos que procesan el plano de usuario (e.g., pesos del *scheduling*, umbrales del control de admisión, configuración de los parámetros de capa de enlace y capa física, etc.). Dichos parámetros podrían haber sido pre-configurados en el equipo por el fabricante en cuestión o bien por el propio operador de la red. Dada la gran flexibilidad que ofrece este esquema, 3GPP ha especificado el comportamiento esperable para un conjunto de QCIs de forma que pueda utilizarse como guía en la configuración de los mecanismos que afectan a la QoS en cada nodo. La finalidad última en la estandarización de un conjunto de QCIs es la de facilitar la consecución de un determinado comportamiento en redes con equipos de diferentes fa-

⁵ La concepción del parámetro QCI es similar al concepto Per Hop Behaviour (PHB) definido en la arquitectura DiffServ para redes IP. En una red DiffServ, la red IP proporciona un conjunto de comportamientos de transferencia diferenciados (i.e., PHBs) tales como Expedited Forwarding (EF) Assured Forwarding (AF) y Best Effort (BE). Los enrutadores de la red a través de los que fluye el tráfico saben que PHB debe aplicarse a cada paquete mediante una etiqueta denominada DSCP (DiffServ Code Point) que se señala en la cabecera del protocolo IP [2].

bricantes así como en escenarios de itinerancia. Los QCI estandarizados se proporcionan en términos de tipo de servicio portador, retardo, tasa de pérdida de paquetes en situaciones de no congestión y un nivel de prioridad [1]. A modo de ejemplo, en la Tabla 3.1 se proporcionan los detalles de algunos de los valores de QCI estandarizados. Los valores de QCI estandarizados son aplicables cualquiera que sea la red de acceso.

Tabla 3.1 Valores de QCI estandarizados

QCI	Tipo de recurso	Prioridad ¹⁾	R e t a r d o Paquete ²⁾	Tasa de pérdida de paquetes ³⁾	Ejemplo de servicio
1	GBR	2	100ms	10 ⁻²	Voz
2	GBR	4	150ms	10 ⁻³	Videoconferencia
5	Non-GBR	1	100ms	10 ⁻⁶	Señalización IMS
6	Non-GBR	6	300ms	10 ⁻⁶	Videostreaming

¹⁾ El valor de prioridad puede ser utilizado por el *scheduler* para ordenar la asignación de recursos entre un conjunto de servicios portadores radio en condiciones tales que, por ejemplo, exista congestión y no se pueda satisfacer el retardo objetivo fijado en ninguno de dichos servicios portadores.

²⁾ Parámetro que se utiliza para determinar el comportamiento del *scheduler* así como la configuración de la capa de enlace (e.g., punto de operación de los mecanismos de retransmisión). El parámetro se plantea como una cota de retardo máximo que deben satisfacer el 98% de los paquetes.

³⁾La tasa de pérdidas se utiliza fundamentalmente para determinar la configuración de la capa de enlace (e.g., punto de operación de los mecanismos de retransmisión RLC y H-ARQ de E-UTRAN). El parámetro representa una cota de la tasa de pérdidas que no son debidas a congestión.

El parámetro ARP se utiliza como un indicador de prioridad en los procesos de establecimiento/modificación/desactivación de un servicio portador. El sistema LTE soporta un total de 15 prioridades. El valor de prioridad puede utilizarse, por ejemplo, en la función de control de admisión cuando no haya suficientes recursos disponibles para dar respuesta a varias activaciones. Asimismo, el valor de ARP también puede utilizarse para desactivar determinados servicios portadores en aras a liberar recursos que deban ser destinados a servicios portadores más prioritarios (*pre-emption*). Esta práctica puede formar parte de un mecanismo de control de congestión.

Para los servicios portadores EPS de tasa garantizada, el parámetro GBR indica justamente la tasa en bits/s que debe proporcionar el servicio portador. Por otro lado, el parámetro MBR acota su tasa máxima de forma que, a través de un mecanismo de control de tasa (rate control), el volumen de tráfico que excede el valor de MBR puede ser descartado. Actualmente, la versión inicial

del sistema LTE solamente proporciona soporte para configuraciones donde los valores de GBR y MBR coincidan (las funciones de control de tasa únicamente se han definido para acotar tasas de pico). En futuras revisiones se espera añadir los mecanismos necesarios para poder implementar un servicio portador de tasa garantizada donde la tasa media sea inferior a la tasa de pico.

Por definición, un servicio portador de tasa garantizada que curse una tasa de datos inferior o igual al valor del parámetro GBR que tenga asociado no tiene que experimentar pérdidas de paquetes por congestión. Por ello, una característica importante de los servicios de tasa garantizada es que deben someterse siempre a control de admisión, ya que su activación conlleva la reserva de un determinado volumen de recursos de transmisión en aras a poder garantizar dicha tasa. Por el contrario, en el caso de los servicios portadores sin tasa garantizada, no resulta estrictamente necesario pasar un control de admisión. Por tanto, a diferencia de un servicio de tasa garantizada, un servicio sin tasa garantizada puede experimentar pérdida de paquetes en situaciones de congestión.

Además de los parámetros de QoS asignados a cada servicio portador EPS, un usuario del sistema LTE tiene asociados dos parámetros adicionales: UE-AMBR y APN-AMBR. Ambos parámetros indican la máxima tasa de transferencia en bits/s que de forma agregada podrán experimentar el conjunto de servicios portadores EPS sin tasa garantizada que tenga activados un usuario. En particular, el parámetro UE-AMBR acota la tasa máxima del equipo de usuario y el parámetro APN-AMBR la tasa agregada máxima del equipo de usuario con una determinada red externa (asociada a dicho APN). Estos dos parámetros forman parte del perfil de subscripción del usuario de forma que, a través de ellos, el operador de la red puede plantear diferentes estrategias de negocio basadas en la tasa máxima de transferencia ofrecida a los usuarios.

El modelo de QoS definido para el sistema LTE simplifica en gran medida el modelo utilizado en UMTS que se basa en un número mucho más elevado de parámetros para indicar el comportamiento de QoS⁶. El servicio portador definido en UMTS proporciona, de forma teórica, mucha más resolución en cuanto a la capacidad de “especificar” un determinado trato de QoS. No obstante, en aplicaciones prácticas, muchos parámetros de QoS en UMTS toman valores por defecto o simplemente no son tenidos en cuenta en el funcionamiento de las funciones de gestión de recursos radio. Asimismo, con respecto a las redes UMTS, en el sistema LTE se ha eliminado la necesidad de tener que incluir en la subscripción del usuario un perfil de QoS autorizado “máximo” para cada uno de los múltiples servicios portadores dedicados que puede activar de forma simultánea un terminal (el perfil de QoS “máximo” serviría para decidir si la QoS solicitada en el establecimiento de cada uno de ellos debe o no ser autorizada). Así, en LTE, el perfil de subscripción de un usuario únicamente contiene el perfil de QoS de los servicios portadores EPS por defecto en cada conexión PDN y los valores APN-AMBR y UE-AMBR comentados anteriormente. Sobre este planteamiento, el perfil de QoS autorizado para los servicios portadores EPS dedicados se obtiene directamente a través del sistema de control de políticas de uso de la red (PCC), que pasa a ser el responsable de determinar dicho comportamiento de QoS requerido en base al tipo de servicio final solicitado y la posible información de subscripción del usuario asociada a la

⁶ Un sistema portador UMTS incluye un total de 4 clases de servicio y 13 atributos de QoS [3].

plataforma de provisión de servicios (y no a la red de acceso). Los detalles de los mecanismos de control de QoS de los servicios portadores mediante el sistema de control de políticas de uso se aborda más detalladamente en el siguiente apartado.

3.2.4 Control del servicio de conectividad: sistema PCC

En los apartados anteriores se han analizado las herramientas de las que dispone el sistema LTE para proporcionar el servicio de conectividad IP con un determinado comportamiento de QoS: conexiones PDN, servicios portadores EPS y parámetros de QoS del sistema. En este apartado se describe el modelo funcional adoptado para gestionar dicho servicio de conectividad.

3GPP ha especificado un subsistema que permite controlar el servicio de conectividad IP ofrecido por la red LTE. Dicho subsistema de control recibe el nombre de PCC (*Policy and Charging Control*) y está concebido para cubrir dos funciones básicas:

- Control mediante políticas de uso (*policy control*⁷) de los servicios portadores EPS. En particular, el control mediante políticas de uso gestionado desde el sistema PCC se centra en dos aspectos concretos: control de los parámetros de **QoS** del servicio portador EPS (denominado como *QoS control* en las especificaciones 3GPP) y control de paso del tráfico a través de la pasarela P-GW (denominado como *gating control* en las especificaciones). Ambos aspectos se desarrollan en detalle más adelante.
- Control de la tarificación del servicio de conectividad, incluyendo el control de servicios pre-pago que requieren un sistema de verificación de crédito *on-line*.

A través del subsistema PCC se articula la interacción del servicio de conectividad proporcionado por LTE con las plataformas que sustentan los servicios finales (e.g., IMS). De esta forma, los servicios portadores EPS que se establezcan en la red LTE están en consonancia con las necesidades de transmisión de los servicios finales cuya señalización se lleva a cabo de forma transparente a través de la red LTE. La determinación de las necesidades de QoS a través del sistema PCC hace que no sea necesario que sea el terminal quien comunique a la red el trato de QoS que espera recibir de ella. Este planteamiento es opuesto al que fue adoptado inicialmente en GPRS donde las características de QoS de los servicios portadores (i.e., contextos PDP) se señalizan desde el equipo de usuario en el momento de su activación. Esta opción se conoce como modelo de QoS iniciado por el móvil y conlleva la necesidad de que en el terminal exista una API (*Application Program Interface*) a través de la cual las aplicaciones puedan solicitar QoS. En cambio, el modelo de QoS propiciado por el subsistema PCC se conoce como modelo de QoS iniciado por la red de forma que, a través de mecanismos de la propia red, ésta puede conocer las necesidades de QoS de las aplicaciones y establecer los servicios portadores adecuados. La posibilidad de activar servicios portadores desde la red

⁷ El término “policy control” utilizado en las especificaciones 3GPP se ha traducido en este libro como “control de políticas de uso”. Es importante matizar que el significado de “policy control” es bastante genérico y puede utilizarse para denotar cualquier mecanismo o sistema de control que utilice reglas o políticas de uso para determinar la configuración u operación de alguna funcionalidad de un sistema (seguridad, movilidad, calidad de servicio, etc.). En el caso de LTE, “policy control” se utiliza para indicar de forma específica “control de QoS” y “control de paso”.

con los parámetros de QoS apropiados (y controlados por el operador) fue una de las mejoras introducidas en las últimas versiones del servicio GPRS en redes UMTS y se ha convertido en el modelo fundamental utilizado en la red LTE.

El subsistema PCC no es específico de la red LTE sino que constituye un subsistema de control transversal que permite gestionar también el servicio de conectividad IP ofrecido por otros tipos de redes de acceso. Por este motivo, en las especificaciones del subsistema PCC se habla de forma genérica de redes IP-CAN (IP *Connectivity Access Network*). Ejemplos de redes IP-CAN son las redes troncales EPC y GPRS, junto con las diferentes redes de acceso radio que admiten. Además, la arquitectura PCC contempla que una determinada red IP-CAN soporte diferentes tecnologías de acceso cuyas características pueden ser tenidas en cuenta a la hora de gestionar las sesiones. En el caso concreto de la red troncal EPC, la red de acceso puede ser E-UTRAN o cualquiera de las otras redes de acceso 3GPP y no 3GPP soportadas, de forma que el subsistema PCC dispone de mecanismos para conocer la red de acceso en uso y puede actuar en consecuencia (e.g., no autorizar el establecimiento de la componente de video de un servicio IMS cuando el terminal no se encuentre en E-UTRAN).

Dada su transversalidad, el subsistema PCC denomina sesión IP-CAN a la asociación formada por una dirección IP, una identidad de usuario (si está disponible) y un identificador de red PDN (e.g., APN). En el caso de la red LTE el concepto de sesión IP-CAN es equivalente al de conexión PDN comentado en el apartado 3.2.1. Asimismo, el tráfico que fluye asociado a una misma sesión IP-CAN, el subsistema PCC puede disociarlo en múltiples flujos de paquetes IP, denominados como *Service Data Flows* (SDFs). Un SDF constituye la unidad mínima de agregación de tráfico sobre la que se aplican las políticas de uso y tarificación del sistema PCC. A diferencia de la asociación directa que existe entre sesión IP-CAN y conexión PDN en una red LTE, es importante matizar que no existe una correspondencia directa entre los SDFs de una sesión IP-CAN y los servicios portadores EPS de la conexión PDN, sino que, cada servicio portador EPS puede transportar uno o múltiples SDFs (e.g., SDFs que tengan la misma QoS pueden transportarse en un único servicio portador EPS). En la Figura 3.5 se ilustra la relación entre los conceptos IP-CAN, SDF, conexión PDN y servicios portadores EPS.

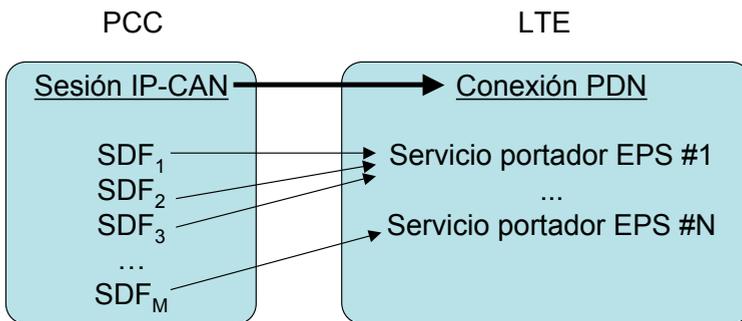


Figura 3.5 Relación entre los conceptos IP-CAN y flujos SDF utilizados en el subsistema PCC con los componentes del servicio de conectividad IP de la red LTE

3.2.4.1 Entidades de red e interfaces

Las entidades funcionales e interfaces de la arquitectura PCC se ilustran en la Figura 3.6 en el contexto de la red LTE. La especificación funcional de dicha arquitectura se realiza en el documento 3GPP TS 23.203 [1].

El elemento central del sistema PCC es la entidad PCRF (*Policy and Charging Rules Function*). Esta entidad es la que determina las políticas de uso (PCC rules) que controlan las sesiones IP-CAN. Las políticas de uso indican, entre otros, el perfil de QoS autorizado (*policy control*) y los parámetros necesarios para el control de la tarificación (*charging control*). Más detalles sobre la información contenida en las políticas de uso se proporcionan posteriormente en el apartado 3.2.4.3. La entidad PCRF se introdujo ya en la UMTS Release 7 como resultado de la integración de las funciones de *Policy Decision Function (PDF)* para gestión y distribución de reglas de QoS y *Charging Function (CRF)* para gestión y distribución de reglas de tarificación que habían sido consideradas en anteriores versiones de UMTS.

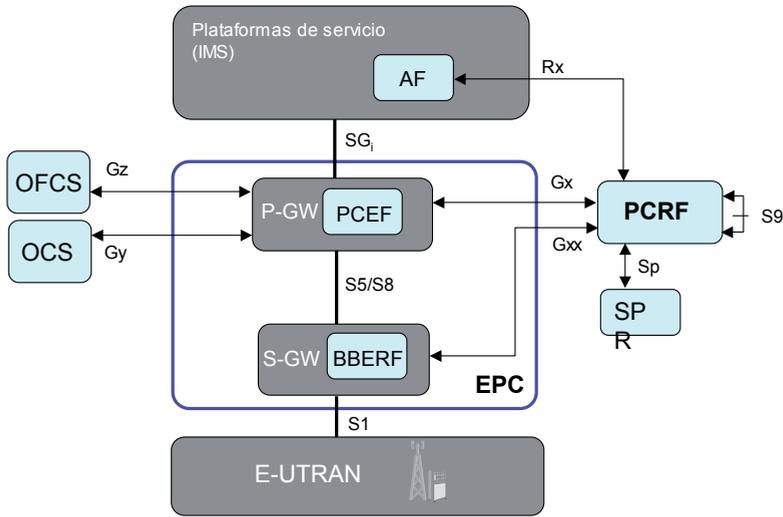


Figura 3.6 Arquitectura del sistema PCC en el contexto de la red LTE

La entidad AF (*Application Function*) es una entidad genérica que representaría a cualquier entidad de la plataforma de los servicios que interacciona con la PCRF mediante la interfaz Rx. En el caso de servicios IMS, las funciones asociadas a AF forman parte de la entidad funcional P-CSCF descrita en el apartado 2.7 del Capítulo 2. La entidad AF es la encargada de extraer la información de sesión necesaria para el uso de PCC a partir de la señalización a nivel de aplicación (e.g., señalización SIP). La entidad AF también puede suscribirse a eventos del sistema PCC relacionados con la provisión de los servicios de transmisión por parte de la red troncal EPC al equipo terminal (e.g., finalización de una conexión PDN, cambio de tecnología de acceso, etc.). Los eventos pueden generarse en las entidades PCRF y BBERF comentadas

a continuación. La comunicación de tales eventos a la entidad AF la realiza la entidad PCRF a través de la interfaz Rx.

La entidad PCEF (*Policy and Charging Enforcement Function*) es la que aplica las reglas PCC recibidas desde la entidad PCRF a través de la interfaz Gx. Esta entidad se ubica en la pasarela entre la red IP-CAN y la red externa. Por tanto, en el caso de la red LTE, las funciones de la entidad PCEF se alojarían en la pasarela P-GW, mientras que en GPRS lo harían en la pasarela denominada como GGSN. La entidad PCEF también proporciona información al PCRF relativa al usuario y a la red de acceso utilizada (e.g., E-UTRAN o accesos alternativos a la red troncal EPC); se encarga de realizar medidas del tráfico cursado (e.g., volumen de tráfico y duración de las sesiones) y de enviar la información relativa al uso de los recursos a las entidades a cargo de las funciones de tarificación (OFCS, *Offline Charging System*, y OCS, *Online Charging System*, introducidas en el apartado 2.6.1 del Capítulo 2).

La entidad BBERF (*Bearer Binding and Event Reporting Function*) únicamente es necesaria en caso de que la gestión de los servicios portadores EPS no se realice desde el P-GW. Esta situación ocurre cuando la interfaz utilizada entre las pasarelas S-GW y P-GW se basa en PMIPv6 o bien cuando se accede al P-GW mediante las interfaces S2a/b/c contempladas en la interconexión de LTE con otras redes de acceso no 3GPP. En estos dos casos, la gestión de los servicios portadores EPS no se realiza desde el P-GW ya que PMIPv6 y el resto de protocolos utilizados en las interfaces S2a/b/c no disponen de los mecanismos de señalización asociados a la gestión de los servicios portadores EPS (nótese que la gestión de servicios portadores EPS en la red troncal se sustenta a través del protocolo GTP-C visto en el apartado 2.6.3.1 del Capítulo 2). En estas situaciones, la entidad BBERF abarca el subconjunto de funciones de la entidad PCEF relacionadas con la asociación de SDFs a servicios portadores EPS y la aplicación de la parte de las reglas PCC concerniente al control de QoS y "gating". El resto de funciones relacionadas con la tarificación siguen realizándose en la P-GW. La comunicación entre las funciones BBERF y la entidad PCRF se realiza a través de la interfaz Gxx.

Finalmente, en la Figura 3.6 se representa también la existencia de una entidad de red denominada SPR (*Subscription Profile Repository*) que constituye una base de datos donde se almacena la información relativa a las políticas de uso de la red que contempla la suscripción de un usuario. A modo de ejemplo, la base de datos SPR puede indicar qué servicios (finales) tiene autorizados un usuario en una determinada red APN, los parámetros de QoS autorizados por servicio, una categoría de usuario (e.g., *business/consumer*), etc. La base de datos SPR se concibe como una entidad funcional diferente de la base de datos de general HSS y, en la versión inicial de las especificaciones del sistema LTE (*Release 8*) no se contempla ninguna interacción entre ambas. En cualquier caso, la integración de las funciones HSS y SPR en un mismo equipo físico podría constituir una posible opción de implementación. La interacción con la entidad SPR se lleva a cabo mediante la interfaz Sp.

Todas las interfaces mencionadas en este apartado (Rx, Gx, Gxx, Sp, Gz, Gy) son aplicaciones específicas del protocolo Diameter. En el apartado 2.6.3 del Capítulo 2 se proporcionan más detalles sobre la utilización de este protocolo en LTE.

3.2.4.2 Marco de funcionamiento

En la Figura 3.7 se ilustra el marco de funcionamiento del subsistema PCC en el contexto de una red LTE. Tal como se ha comentado en el apartado 3.2.1, el equipo de usuario LTE puede acceder a servicios proporcionados por una plataforma como IMS a través del establecimiento de una conexión PDN y del servicio portador EPS por defecto asociado. La señalización a nivel de aplicación entre el equipo de usuario y la plataforma de servicios fluye a través del servicio portador EPS por defecto (paso (1) mostrado en la Figura 3.7). Dicha señalización es procesada por la entidad AF de la plataforma de servicios, que en el caso de IMS forma parte de las funciones del P-CSCF (descrito en el apartado 2.7 del Capítulo 2). A partir de la señalización a nivel de aplicación, la entidad AF determina las características del servicio al que accede el usuario y las comunica a la entidad PCRF del subsistema PCC para que genere las políticas de uso y tarificación correspondientes (paso (2) en la Figura 3.7). Además de la información proporcionada por la entidad AF, la determinación de las políticas de uso en la entidad PCRF puede basarse en las siguientes fuentes de información:

- Información asociada a la suscripción del usuario en relación con las estrategias de uso de los servicios de la red (e.g., un usuario está autorizado a activar sesiones que transmitan video con un códec que requiere una determinada tasa de transferencia garantizada). Este tipo de información se almacena en la base de datos SPR. El acceso a dicha información se realiza a través de la interfaz Sp, tal como se muestra en el paso (3) de la Figura 3.7)
- Configuraciones pre-establecidas en el PCRF por el propio operador de las políticas de uso aplicables a determinados servicios.
- Información de las características de las redes de acceso utilizadas obtenidas a través de las interfaces Gx y/o Gxx (e.g., la entidad PCRF conoce si el terminal está accediendo a través de E-UTRAN o bien de otra red de acceso radio alternativa).

Finalmente, tal como se ilustra en el paso (4) de Figura 3.7, las políticas de uso y tarificación generadas se envían a las entidades de red (e.g., PCEF) encargadas de su aplicación en el plano de usuario.

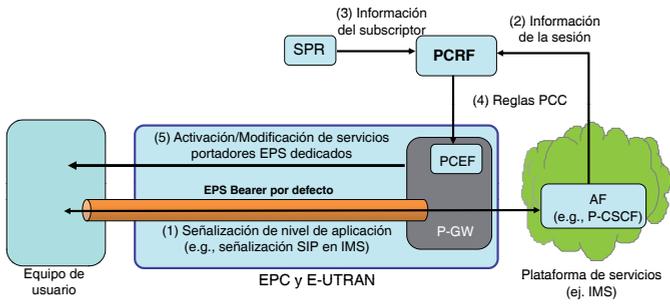


Figura 3.7 Funcionamiento básico del modelo PCC

3.2.4.3 Políticas de uso

El contenido básico de una regla de uso (regla PCC) se muestra en la Figura 3.8. Cada regla de uso, identificada mediante un campo específico, contiene una descripción del flujo de datos SDF al que aplica la regla junto con los datos asociados al control mediante políticas de uso (*policy control*) y al control de tarificación.

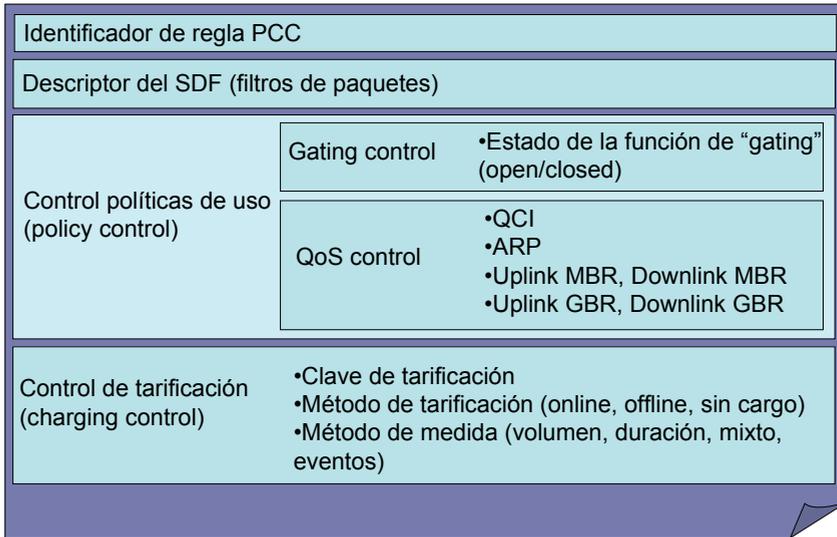


Figura 3.8 Ilustración del contenido de una regla PCC

Los parámetros utilizados para el control mediante políticas de uso se dividen en dos bloques:

- Parámetro de control de la función de paso (*gating control*). Mediante este parámetro, se controla el flujo de tráfico del SDF a través de la pasarela P-GW (o S-GW, en el caso de utilizar PMIPv6 en S5/S8). Así, desde la entidad PCRF se puede bloquear o dejar pasar el tráfico IP de un usuario asociado a un determinado servicio.
- Parámetros de control de la calidad de servicio. En la regla PCC se especifican los cuatro parámetros descritos en el apartado 3.2.3: QCI, ARP, MBR y GBR, los dos últimos sólo presentes en el caso de requerirse una tasa de transferencia garantizada. Los parámetros de QoS de la regla PCC aplican a un SDF.

La regla PCC también incluye los parámetros que controlan la tarificación del servicio. Así, en la regla PCC se incluyen:

- Clave de tarificación. Este parámetro es un puntero al tipo de tarifa que se aplicará en el sistema OFCS o OCS para este flujo de datos (e.g., tarifa plana).
- Método de tarificación. Indica si la tarificación del SDF debe realizarse *online* (e.g., me-

diante el acceso en tiempo real a un sistema de control de crédito), *offline* (una vez terminado el servicio) o bien no está sujeto a tarificación.

- Método de medida. Mediante este parámetro se indica el tipo de información que debe obtenerse para tarificar el servicio. Dicha información puede ser volumen de datos transferidos, duración del servicio, eventos asociados al servicio, y combinaciones de las anteriores.

3.2.5 Procedimientos de gestión de sesiones

En el sistema LTE la activación/modificación/desactivación de los servicios portadores EPS se controla desde la propia red en base a los datos de subscripción del usuario y/o a las políticas de uso recibidas desde el sistema PCC. Tal como se ha comentado en el apartado 3.2.4, es importante destacar que este planteamiento es diferente al seguido en el servicio GPRS de redes UMTS donde la activación de los servicios portadores la inicia el terminal.

Los procedimientos principales relacionados con la gestión de sesiones son:

- Procedimiento de registro (*Network Attach*). A través de este procedimiento se establece el servicio de conectividad IP que ofrece la red LTE. Existen diferentes variantes del procedimiento de registro en función de si la red de acceso utilizada es E-UTRAN o cualquiera de las otras redes de acceso alternativas 3GPP y no 3GPP contempladas.
- Procedimiento de petición de servicio (*Service Request*). El modelo de servicio ofrecido por la red LTE permite que un usuario en modo *idle* (i.e. sin una conexión a E-UTRAN) mantenga abiertos los servicios portadores EPS en la red troncal. Este procedimiento permite una re-activación rápida del plano de usuario cuando el terminal pasa de *idle* a conectado (los estados de movilidad se analizarán más detalladamente en el apartado 3.3 de este mismo capítulo).
- Petición de conexión PDN solicitada por el terminal (*UE Requested PDN Connectivity*). El sistema LTE permite el equipo de usuario inicie el procedimiento de establecer una conexión PDN adicional a la conexión PDN establecida en el procedimiento de registro.
- Activación, modificación y desactivación de los servicios portadores EPS dedicados (*EPS Bearer Activation/Modification/Deactivation*). La gestión de los servicios portadores EPS dedicados es uno de los pilares de la gestión de sesiones en la red LTE. La activación y modificación de estos servicios puede estar vinculada al control dinámico de QoS ofrecido por el subsistema PCC.
- Modificación del servicio portador solicitada por el terminal (*UE requested bearer resource modification*). Este procedimiento permite que el terminal pueda solicitar cambios en los servicios portadores que le ofrece la red. Los cambios pueden ser tanto un cambio de los parámetros de QoS como una modificación de los filtros de paquetes que determina la composición del tráfico agregado en un servicio portador. El procedi-

miento constituye una vía de escape al modelo general de activación de servicios portadores donde, tal como se ha comentado al principio de este apartado, es la red quien controla su activación y determina sus características. Si la solicitud de modificación realizada por un terminal es aceptada por la red, ésta procede a iniciar los mecanismos pertinentes de activación, modificación y/o desactivación de los servicios portadores EPS (el control sigue teniéndolo la red, pero en este caso, atendiendo a una petición proveniente del terminal).

A continuación se describen más detalladamente el procedimiento de registro, el de activación de un servicio portador EPS dedicado y el procedimiento de petición de servicio cuando el terminal se encuentra en modo *idle*. La descripción de estos tres procedimientos básicos proporciona una visión bastante completa del alcance de la gestión de sesiones en el sistema LTE. El lector interesado en profundizar más en la operativa de estos procedimientos, o bien en el resto de procedimientos de gestión de sesiones expuestos anteriormente y no abordados de forma explícita en el libro, puede encontrar su descripción en la especificación 3GPP TS 23.401 [4].

3.2.5.1 Procedimiento de registro

El procedimiento de registro es el primer procedimiento que ejecuta un usuario del sistema LTE en aras a poder recibir los servicios de la red. El procedimiento de registro normalmente se lleva a cabo cuando se enciende el equipo de usuario y éste detecta la presencia de una red LTE. A diferencia de sus predecesores GSM y UMTS donde los procedimientos de registro correspondientes (*IMSI Attach* y *GPRS Attach*) están asociados exclusivamente a la gestión de movilidad, en LTE dicho procedimiento también forma parte de la gestión de sesiones. El motivo radica en que el procedimiento de registro en LTE conlleva el establecimiento de una conexión PDN a través de la activación de un servicio portador EPS por defecto y, opcionalmente, servicios portadores EPS dedicados adicionales. Por tanto, en LTE, una vez el terminal ya se ha registrado en la red, ya dispone de un servicio de conectividad IP operativo (faceta popularmente conocida como “*always-on*”). En la Figura 3.9 se ilustra un procedimiento de registro, cuyos detalles se comentan a continuación.

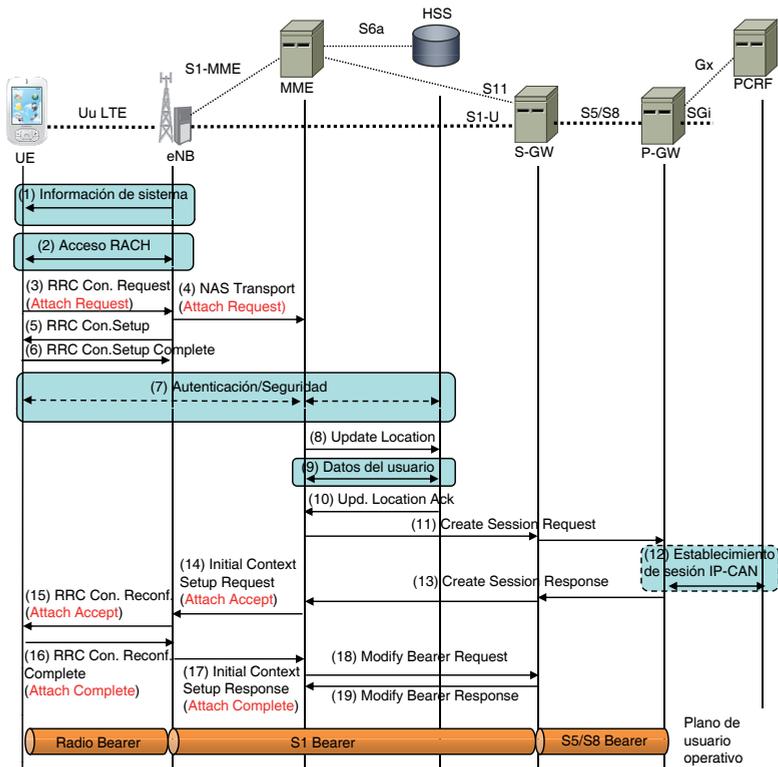


Figura 3.9 Procedimiento de registro

Los pasos identificados en la Figura 3.9 son los siguientes:

(1) Información de sistema

Previamente al inicio del procedimiento de registro, el terminal LTE decodifica la información de sistema que recibe de la red LTE. Esta información se difunde a través de los canales *broadcast* que soporta la interfaz radio (i.e., canal BCCH descrito en el Capítulo 5). La información de sistema permite que los terminales conozcan los parámetros de operación de la red necesarios para proceder a la realización del registro. En particular, entre la información de sistema recibida relacionada con el procedimiento de registro, cabe citar:

- Identidad de la celda que está decodificando el terminal.
- Identidad del área de seguimiento (*Tracking Area Identifier, TAI*) a la que pertenece la celda. La definición y operativa de las áreas de seguimiento se analiza más detalladamente en el apartado 3.3.2 de este mismo capítulo.
- Una lista de las redes móviles (i.e., lista de PLMNs) accesibles a través de la celda (esta opción posibilita que una misma estación base LTE pueda proporcionar servicio a múltiples operadores de red, i.e., *network sharing*).

- Permisos de acceso a la celda. Mediante esta opción el operador de la red puede restringir el acceso a la celda a determinados colectivos de usuarios (e.g., se han creado 16 clases de acceso entre las que se encuentran usuarios convencionales, servicios de emergencia, operarios de la red, etc.).
- Criterios radio para la selección de la celda. Cada celda difunde un conjunto de parámetros (e.g., niveles de potencia y calidad de recepción) que utilizan los terminales en los procesos de selección y re-selección de celda [6] para determinar cuál es la celda a la que deben estar sintonizados cuando se encuentran en estado *idle*.

(2) Acceso aleatorio

En base a la información anterior, el terminal LTE decide la celda a través de la cuál debe iniciar el procedimiento de registro. En dicha celda, lo primero que debe hacer el terminal es establecer una conexión de control RRC. Para ello, el terminal inicia el procedimiento de acceso aleatorio a través del canal RACH de la interfaz radio. Los detalles de la operativa de acceso a través del canal RACH se proporcionan en el Capítulo 5.

(3) Inicio de la conexión RRC y mensaje de petición de registro

A raíz de la ejecución del procedimiento de acceso aleatorio, el terminal envía el mensaje RRC de petición de establecimiento de conexión. Dicho mensaje, además de contener los campos específicos del protocolo RRC, transporta en su parte de datos el mensaje de señalización NAS correspondiente a la petición de registro (mensaje "Attach Request"). El mensaje NAS de petición de registro, junto con la señalización NAS que se deriva asociada de esta petición, se muestra en color rojo en la Figura 3.9 para facilitar su seguimiento. El mensaje de petición de registro forma parte del protocolo EMM descrito en el apartado 2.6.3.5 del Capítulo 2, en el contexto de la señalización NAS entre equipo de usuario y red troncal EPC.

El mensaje de petición de registro identifica al usuario mediante su IMSI o una identidad temporal denominada GUTI⁸ (Globally Unique Temporary Identity) que el usuario tenga guardada como resultado de un registro previo. Junto con la identificación del usuario, el mensaje contiene también información sobre la última área de seguimiento donde estuvo registrado el terminal, el tipo de conexión PDN que solicita (e.g., IPv4, IPv6 o ambas, así como, opcionalmente, la identidad APN de la red externa), las capacidades del terminal (e.g., modos de recepción discontinua soportados, parámetros de seguridad), etc.

(4) Envío del mensaje de petición de registro a la entidad MME

El mensaje RRC que recibe el eNB contiene campos que indican también la identidad de red PLMN en la que el usuario quiere registrarse (en caso de que haya múltiples opciones) y

⁸ El identificador GUTI es un identificador único que asigna la entidad MME a un usuario en el momento de registro o actualización del área de seguimiento. El GUTI se compone de un identificador de usuario único en el contexto de un MME (denominado TMSI, Temporary Mobile Subscriber Identity) y de un identificador universal de la entidad MME (denominado GUMMEI, Globally Unique MME Identifier). Mediante el GUTI, un usuario dispone de un alias con validez universal, a diferencia de los alias utilizados en UMTS y GSM (equivalentes al identificador TMSI integrado en el GUTI). La relación de identificadores utilizados en LTE se especifica en el documento TS 23.003 [5].

la identidad de la entidad MME (i.e., GUMMEI⁹) de esta red PLMN en la que estuvo previamente registrado el terminal (si la hubo). Esta información permite que el eNB envíe el mensaje “*Attach Request*” a la entidad MME apropiada, teniendo en cuenta también que un eNB puede proporcionar acceso a múltiples entidades MME de la misma red PLMN. El mensaje “*Attach Request*” se envía a través de la interfaz S1-MME que, como se vio en el Capítulo 2, soporta el envío transparente de mensajes NAS al igual que el protocolo RRC en la interfaz radio. En caso de que la entidad MME indicada por el usuario no pertenezca al conjunto de MMEs asociadas con el eNB, éste puede escoger una entidad MME alternativa de la misma red PLMN.

(5) y (6) Establecimiento de la conexión RRC

Al tiempo que progresa la señalización asociada al registro en la infraestructura, la conexión RRC con el terminal en la interfaz radio termina estableciéndose de forma que queda activado un servicio portador radio de señalización (i.e., SBR1) a través del cual se soportan los posteriores intercambios de señalización NAS (véase el Capítulo 5 para una descripción más detallada de las características de los servicios portadores de señalización radio).

(7) Autenticación y Seguridad

Una vez recibido el mensaje de “*Attach Request*” y conocida la identidad del usuario correspondiente en la entidad MME, puede llevarse a cabo el mecanismo de autenticación y seguridad descrito posteriormente en el apartado 3.4.4.1 de este mismo capítulo. El mecanismo de autenticación y seguridad es necesario siempre que en la entidad MME no exista un contexto del usuario previamente establecido (o recuperado de la entidad MME previa en la que el usuario en cuestión estuviera registrado) con una asociación de seguridad válida. Si existe dicha información, el procedimiento de autenticación puede omitirse.

(8), (9) y (10) Actualización de la localización

Si la entidad MME ha cambiado desde el último registro o si ésta no dispone de la información de suscripción del usuario, la entidad MME inicia una actualización de la localización en la base de datos HSS del sistema. Este proceso permite que la HSS siempre almacene información de dónde se encuentran accesibles los usuarios registrados. Además, como resultado del procedimiento de actualización de la localización, la entidad MME recibe los datos de suscripción del usuario que indican la información sobre las posibles conexiones PDN autorizadas al usuario junto con el perfil de QoS suscrito (i.e., QCI, ARP, APN-AMBR para configurar el servicio portador EPS por defecto en cada PDN, además del parámetro UE-AMBR). Toda esta información se guarda en la entidad MME en una estructura de datos que se denomina como contexto. El contenido completo de contexto de un usuario en la entidad MME se proporciona en 3GPP TS 23.401 [4]. La entidad MME mantiene un contexto activo para cada uno de los usuarios que se registren a través de ella.

⁹ Globally Unique MME Identity (GUMMEI): identidad universal que identifica a una entidad MME. Forma parte del identificador GUTI asignado al terminal en procedimiento de registro.

(11), (12) y (13) Establecimiento de la conexión PDN

La entidad MME inicia el establecimiento de la conexión PDN con la red externa correspondiente que resultará en la activación de, como mínimo, el servicio portador EPS por defecto.

La selección de la red externa y de la pasarela P-GW correspondiente puede basarse en la identidad APN enviada por el usuario en el mensaje de registro y en los parámetros de suscripción, si dicha información está presente. A nivel de la suscripción, además de poderse indicar la identidad APN por defecto, también es posible asociar una dirección IP estática a un usuario e incluso la identidad de la pasarela P-GW (si hubiera varias) que debe proporcionarle acceso a la red externa. En caso de que no fuera posible determinar la red externa de la información de suscripción, y el usuario tampoco indicara ningún identificador APN, la entidad MME escogería la red PDN que tuviera configurada por defecto.

Una vez seleccionada la red externa y la pasarela P-GW asociada en la que iniciar la conexión PDN, la entidad MME envía un mensaje "*Create Session Request*" hacia la entidad S-GW que servirá como punto de anclaje del plano de usuario de dicha conexión. La elección de la entidad S-GW la realiza la entidad MME atendiendo a criterios tales como balanceo de carga u optimización del rutado dentro de la red LTE. El mensaje "*Create Session Request*" incluye información sobre la identidad del usuario, la identidad del servicio portador EPS que se está estableciendo, sus parámetros de QoS (atendiendo a los datos de suscripción), la dirección de la pasarela P-GW, el tipo de red de acceso a través de la que accede el terminal¹⁰, opciones de configuración del protocolo IP¹¹, etc. Tras la recepción del mensaje "*Create Session Request*", la entidad S-GW crea otro mensaje análogo y lo envía hacia la pasarela P-GW indicada por el MME. El mensaje enviado por el S-GW añade a los parámetros comentados anteriormente la información necesaria para establecer el túnel en el plano de usuario entre S-GW y P-GW (e.g., parámetro TEID del protocolo GTP-U visto en el apartado 2.6.3.2 del Capítulo 2).

Si la dirección IP no está fijada por suscripción, la pasarela P-GW es la encargada de determinar la dirección IP que le será comunicada al terminal mediante la señalización de establecimiento de la conexión PDN. En este caso, la dirección IP seleccionada por la pasarela P-GW podría ser una dirección dinámica (gestionada por la propia pasarela o bien adquirida a través de servidores de direcciones IP, e.g., servidores DHCP) así como una dirección estática (e.g., pre-asignada de forma permanente al usuario y guardada, e.g., en un servidor DHCP, pero no como parámetro de suscripción en el HSS). En el caso de servicios de conectividad IPv4, también existe la posibilidad de terminar el establecimiento de la conexión PDN sin asignarle ninguna dirección IP al terminal. La obtención de la dirección IP se realiza posteriormente mediante señalización DHCPv4 iniciada por el propio terminal.

En caso de que se utilice el subsistema PCC visto en el apartado 3.2.4 para el control del

¹⁰ El establecimiento del servicio portador podría ser consecuencia también del registro de un terminal que accede a través de UTRAN o GERAN, con lo que la información sobre tipo de RAT se incluye entre los parámetros asociados al servicio portador EPS entre S-GW y P-GW.

¹¹ A partir de la información de suscripción, la entidad MME debe indicar a la entidad P-GW si el usuario dispone de una asignación concreta de dirección IP o, en caso contrario, de cual debe ser el mecanismo utilizado para la asignación de la dirección IP (e.g., DHCPv4, etc.).

servicio de conectividad de la red LTE, la pasarela P-GW inicia el procedimiento de establecimiento de sesión IP-CAN con la entidad PCRF correspondiente (véase paso (12) de la Figura 3.9). Tal como se ha comentado en el apartado 3.2.4, una sesión IP-CAN es una asociación entre la red LTE y el sistema PCC en base a una dirección IP, una identidad de usuario (IMSI en el caso de LTE) y un identificador de red PDN (e.g., APN). En la activación de la sesión IP-CAN, la entidad P-GW podría recibir un conjunto de reglas PCC por defecto aplicables al usuario. Dichas reglas pueden conllevar la modificación de los parámetros de QoS recibidos en la pasarela P-GW a raíz de la señalización iniciada desde la entidad MME así como el establecimiento de servicios portadores EPS dedicados además del servicio portador por defecto.

Finalizada la señalización del establecimiento de la conexión PDN (mensaje “*Create Session Response*”), el plano de usuario entre P-GW y S-GW queda operativo y en las tres entidades de la red troncal (MME, S-GW y P-GW) quedan establecidos contextos de datos asociados al usuario donde se guardan las configuraciones de los servicios portadores establecidos.

(14) Creación de un contexto en E-UTRAN y respuesta a la petición de registro

La entidad MME envía al eNB el mensaje de control “*Initial Context Setup Request*” mediante el protocolo S1-AP soportado en la interfaz S1-MME (véase apartado 2.5.3.2 del Capítulo 2). Mediante este mensaje se inicia la creación de un contexto para sustentar la operativa del plano de usuario del terminal en proceso de registro. Este mensaje contiene, entre otros, la identidad del servicio (o servicios) portador EPS activado y sus parámetros de QoS, el parámetro UE-AMBR (que permite al eNB limitar la tasa de transferencia máxima entre el usuario con la red), posibles restricciones de movilidad que tenga el usuario (i.e., lista de restricciones de *handover*), los parámetros de seguridad utilizados en la operativa de la interfaz radio (i.e., la interfaz radio ofrece un servicio de confidencialidad e integridad cuyas claves secretas se derivan a partir de la información enviada desde la entidad MME) y los parámetros necesarios para establecer el plano de usuario entre el eNB y el S-GW (i.e., dirección IP de la pasarela S-GW e identificador de túnel TEID).

En el mensaje de control de la interfaz S1-MME, se transporta también el mensaje “*Attach Accept*” correspondiente a la señalización entre el equipo de usuario y la entidad MME. Este mensaje es la respuesta al mensaje de registro enviado por el terminal en el paso (3) de la Figura 3.9. El mensaje “*Attach Accept*” contiene el nuevo identificador GUTI asignado al usuario y la lista de áreas de seguimiento donde es válido el registro efectuado. El mensaje incluye también la identidad del servicio (o servicios) portador EPS establecido, sus parámetros de QoS, el identificador APN de la red externa y la dirección IP asignada, o en su caso, las indicaciones necesarias para obtenerla. Finalmente, el mensaje también transporta información de seguridad relacionada con los servicios de confidencialidad e integridad de los protocolos NAS así como, en caso de que el terminal soporte interfaces UTRAN o GERAN, identificadores y parámetros de QoS que serán utilizados para sustentar la operativa de los procedimientos de movilidad entre redes de acceso.

(15) y (16) Establecimiento de los servicios portadores en la interfaz radio

El eNB envía el mensaje "*RRC Connection Reconfiguration*" para establecer el servicio (o servicios) portador radio. Tal como se indica en el apartado 3.2.2, existe una correspondencia única entre un servicio portador EPS y un servicio portador radio, de forma que, en caso de establecerse servicios portadores EPS dedicados adicionales al servicio por defecto, la interfaz radio establece el número de servicios portadores radio que corresponda. Más detalles del procedimiento de establecimiento de los servicios portadores radio se proporcionan en el Capítulo 5.

Después del mensaje (16) el plano de usuario en la dirección ascendente ya se encuentra operativo, no así el descendente que todavía requiere el establecimiento del plano de usuario entre el eNB y el S-GW.

(16) y (17) Finalización del registro

Finalmente, el equipo de usuario envía el mensaje que completa el proceso de registro ("*Attach Complete*") que se transporta hasta la entidad MME mediante el servicio de transferencia de mensajes NAS que ofrece el protocolo RRC de la interfaz radio y el protocolo S1-AP de la interfaz S1-MME.

(18) y (19) Finalización de la configuración del servicio portador EPS

El plano de usuario en el enlace descendente queda completamente establecido una vez el S-GW dispone de los parámetros que caracterizan la terminación del túnel GTP en el eNB (i.e., dirección IP del eNB e identificador TEID del túnel GTP). A tal efecto, en el mensaje de respuesta a la creación del contexto ("*Initial Context Setup Response*") que el eNB envía a la entidad MME se incluyen los parámetros para la terminación del túnel del enlace descendente. Posteriormente, y una vez que la entidad MME haya recibido también el mensaje de "*Attach Complete*", la entidad MME inicia el procedimiento de modificación de las características del servicio portador (mensaje "*Modify Bearer Request*") para incorporar los parámetros de terminación del túnel GTP en el contexto asociado al servicio portador del usuario en cuestión en el S-GW.

3.2.5.2 Procedimiento de activación de servicios portadores dedicados

Tal como se ha indicado en el apartado 3.2.2, una conexión PDN puede albergar, además del servicio portador EPS por defecto, múltiples servicios portadores EPS dedicados. El establecimiento de un servicio portador dedicado sigue el diagrama de mensajes representado en la Figura 3.10.

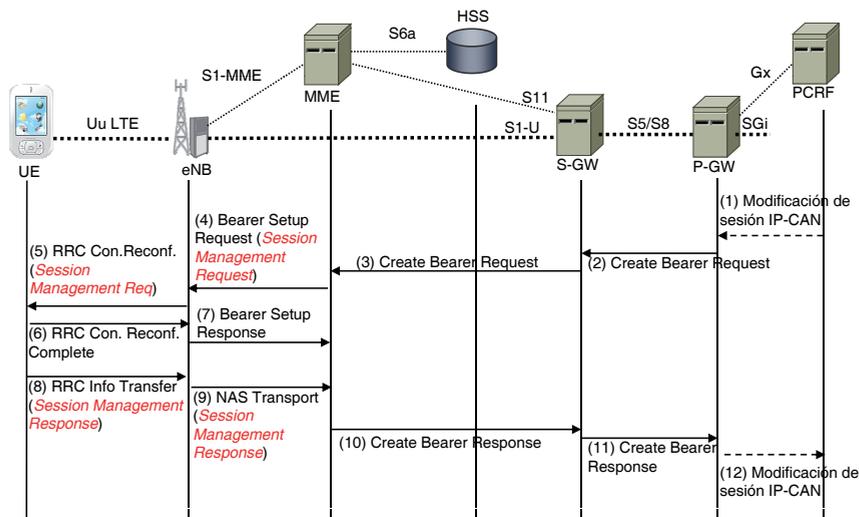


Figura 3.10 Procedimiento establecimiento de un servicio portador EPS dedicado

Los pasos identificados en la Figura 3.10 son los siguientes:

(1) Inicio del procedimiento

En caso de que se emplee el sistema PCC, el inicio del procedimiento de activación del servicio portador dedicado puede originarse como consecuencia de un procedimiento de modificación de la sesión IP-CAN (establecida en el proceso de registro tal como se descrito en el paso (12) de la Figura 3.9). La modificación de la sesión IP-CAN puede ser consecuencia de la interacción del sistema PCC con la plataforma de servicios (tal como se verá más adelante en el ejemplo descrito en la Figura 3.11) así como de la propia interacción de la pasarela P-GW con el sistema PCC (e.g., la pasarela P-GW notifica que el terminal ha cambiado de red de acceso y, en consecuencia, el sistema PCC decide cambiar la configuración del servicio de conectividad). La modificación de la sesión IP-CAN puede conllevar la transferencia de un conjunto de reglas PCC nuevas desde la entidad PCRF que requieran activar un nuevo servicio portador EPS dedicado (también podría darse el caso que las nuevas reglas únicamente requieran la modificación de un servicio portador existente).

Si no se utiliza el sistema PCC, la activación de los servicios portadores dedicados podría iniciarse como resultado de la aplicación de unas reglas PCC guardadas localmente en la pasarela P-GW. En este caso, una vez realizado el registro y establecido el servidor portador EPS por defecto, la propia pasarela P-GW puede iniciar automáticamente el establecimiento de servicios portadores EPS dedicados. A modo de ejemplo, un operador LTE podría ofrecer un servicio de conectividad que, por defecto, activara siempre un servicio portador EPS adicional a través del cual se enviaría el tráfico de una determinado servicio final (el servicio portador dedicado permanecería establecido de forma permanente sin estar vinculado al uso del servicio final en cuestión).

(2) y (3) Petición de creación del servicio portador EPS dedicado

La activación del servicio portador EPS dedicado se inicia desde la pasarela P-GW mediante el envío del mensaje "*Create Bearer Request*" a la pasarela S-GW, en caso de utilizarse la interfaz S5/S8 basada en GTP, tal como sería el caso del ejemplo en la Figura 3.10. El mensaje de activación contiene los parámetros de QoS derivados de las reglas PCC aplicables (i.e., QCI, ARP, GBR y MBR), junto con el descriptor de tráfico TFT asociado (el descriptor de tráfico se deriva de los filtros de paquetes contenidos en las reglas PCC, véase Figura 3.8). El mensaje también contiene la información necesaria para configurar el plano de usuario entre pasarelas. Finalmente, el mensaje de activación del servicio portador dedicado llega a la entidad MME que está a cargo del usuario en cuestión.

(4) Petición de activación del servicio portador radio a E-UTRAN y al equipo de usuario

La entidad MME selecciona un identificador para el nuevo servicio portador y construye un mensaje NAS denominado "*Session Management Request*". Dicho mensaje contiene la identidad del servicio portador y sus parámetros de QoS (excepto el ARP) y el filtro de paquetes TFT asociado para ser aplicado en el equipo de usuario para seleccionar el tráfico ascendente. Asimismo, si el terminal soporta interfaces UTRAN o GERAN, el mensaje incorpora identificadores y parámetros de QoS (e.g., parámetros de QoS de un contexto PDP derivados a partir de los parámetros del servicio portador EPS) que serán utilizados para sustentar la operativa de los procedimientos de movilidad entre redes de acceso. La entidad MME incluye el mensaje "*Session Management Request*" encapsulado en el mensaje de petición de activación del servicio portador radio que envía al eNB ("*Bearer Setup Request*"). En aras a facilitar su seguimiento, en la Figura 3.10 el mensaje NAS "*Session Management Request*" y las respuestas asociadas al mismo se han resaltado en color rojo. El mensaje "*Bearer Setup Request*" contiene los parámetros de QoS del servicio portador (aquí sí que se incluye el ARP, pero no se incluye el TFT) y el identificador de túnel para el establecimiento del plano de usuario entre el eNB y el S-GW en sentido ascendente.

Si en el momento en que la entidad MME recibe el mensaje de "*Create Bearer Request*" de la pasarela S-GW, el usuario se encuentra en modo *idle* (y por tanto no se encuentra conectado a ningún eNB), la entidad MME inicia previamente el procedimiento de petición de servicio ("*Network Triggered Service Request*") que conlleva el aviso del usuario a través de los canales de aviso (i.e. canales PCCH) de los eNBs que forman parte del área de seguimiento donde se encuentra localizado el terminal. El procedimiento de petición de servicio se describe más detalladamente en el apartado 3.2.5.3.

(5) y (6) Establecimiento del servicio portador radio

El eNB mapea los parámetros de QoS del servicio portador EPS a los del servicio portador radio asociado y envía un mensaje RRC para reconfigurar los parámetros de la interfaz radio en el equipo terminal. El mensaje RRC transporta el mensaje "*Session Management Request*" proveniente de la entidad MME. La activación del servicio portador radio finaliza con el envío del mensaje "*RRC Connection Reconfiguration Complete*" por parte del terminal.

(7) Confirmación del establecimiento del servicio portador en E-UTRAN

Una vez establecido el nuevo servicio portador radio, el eNB responde a la petición de activación originada desde la entidad MME. El mensaje contiene los parámetros necesarios para establecer el plano de usuario entre el eNB y S-GW ahora en sentido descendente.

(8) y (9) Confirmación del establecimiento del servicio portador en el equipo de usuario

La confirmación se realiza mediante el envío del mensaje NAS "Session Management Response" que se transporta hasta la entidad MME mediante el protocolo RRC de la interfaz radio y el protocolo S1-AP de la interfaz S1-MME.

(10) y (11) Respuesta a la petición de la pasarela P-GW

Tras la recepción de la respuesta del eNB ("Bearer Setup Response") y la del equipo de usuario ("Session Management Response"), la entidad MME responde a la petición de activación del servicio portador dedicado a la pasarela S-GW (mensaje (10)) y ésta a la pasarela P-GW (mensaje (11)). En dichas respuestas se incluyen los identificadores de túnel que permiten terminar de establecer el plano de usuario entre eNB y P-GW en sentido descendente.

(12) Respuesta a la modificación de la sesión IP-CAN

Si la activación del servicio portador EPS dedicado se inició a partir del envío de reglas PCC desde el PCRF, la pasarela P-GW notifica a la entidad PCRF el resultado de la activación.

Tal como se ha comentado al inicio de la descripción del procedimiento de activación del servicio portador EPS dedicado, la activación de estos portadores puede ser debida a la interacción entre las plataformas de servicios y la red LTE. A modo de ejemplo, en la Figura 3.11 se ilustra cómo el establecimiento de un servicio final a través del subsistema IMS conduce a la activación del servicio portador dedicado en la red LTE.

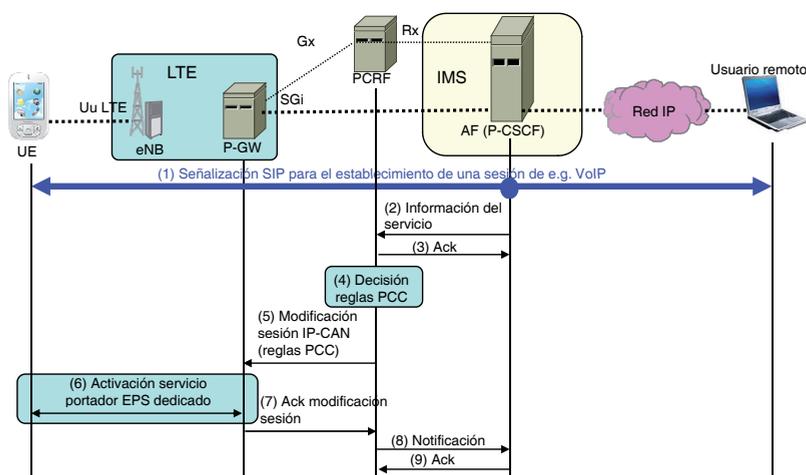


Figura 3.11 Interacción con la plataforma de servicios

Los pasos mostrados en la Figura 3.11 se describen a continuación:

(1) Señalización propia del servicio final

El ejemplo mostrado podría corresponder al establecimiento de una sesión de voz sobre IP (VoIP) a través del subsistema IMS. Tal como se describe en el apartado 2.7 del Capítulo 2, la señalización de nivel de aplicación en IMS se realiza mediante el protocolo SIP. En el caso de un servicio VoIP, los mensajes SIP son intercambiados entre el terminal LTE y el equipo remoto con el que se esté estableciendo la sesión de voz. La señalización SIP fluye a través de los servidores de control CSCF del dominio IMS. En particular, el servidor P-CSCF es el que actuaría como la entidad AF definida en el sistema PCC (véase el apartado 3.2.4.1 donde se describen las entidades del sistema PCC).

(2) y (3) Información de las características del servicio final

Así, a partir de la señalización SIP cursada, el servidor P-CSCF conoce las características del servicio final que se está negociando (e.g., presencia de componentes de audio y/o video y codecs correspondientes) y puede informar a la entidad PCRF de las características del servicio. La información sobre el servicio final puede ir acompañada de un registro para la recepción de notificaciones asociadas a eventos que puedan acaecer en la sesión IP-CAN relacionada que soporta dicho servicio (e.g., notificación de la activación de un servicio portador IP-CAN para sustentar el servicio final).

(4) y (5) Decisión de nuevas reglas PCC

Con la nueva información sobre el servicio, la entidad PCRF decide la necesidad de realizar alguna modificación en la conexión PDN y, en su caso, genera las nuevas reglas PCC que son enviadas a la pasarela P-GW

(6) y (7) Activación de un servicio portador EPS dedicado

El procedimiento de activación es justamente el explicado anteriormente en este mismo apartado (véase Figura 3.10). Una vez finalizada la activación, la pasarela notifica el resultado a la entidad PCRF.

(8) y (9) Notificación a la plataforma de servicios

En caso de que la entidad AF hubiera solicitado la notificación del evento asociado al establecimiento del servicio portador en la comunicación de la información del servicio (mensaje (2)), la entidad PCRF genera la notificación correspondiente.

3.2.5.3 Procedimiento de petición de servicio

Una característica importante de la gestión de sesiones en la red LTE es que un equipo de usuario puede desactivar su conexión radio con la red de acceso (i.e., entrar en modo *idle*) pero mantener establecidos los servicios portadores EPS en la red troncal EPC. De esta forma, el re-establecimiento de la conectividad entre el equipo de usuario y la red externa se

puede realizar de forma muy rápida una vez el terminal vuelve a conectarse a E-UTRAN. El procedimiento utilizado para solicitar la reactivación tanto del plano de control (señalización NAS) como del plano de usuario de un equipo de usuario que se encuentra en modo *idle* se denomina en las especificaciones como *Service Request*. El motivo principal que conduce a la ejecución del procedimiento de petición de servicio es la existencia de tráfico de datos y/o señalización pendiente de ser enviado en el terminal o en la red troncal EPC.. El procedimiento de petición de servicio puede ser iniciado por el propio terminal o bien por la red mediante la ejecución previa de un procedimiento de aviso (*Paging*). Es este caso, el procedimiento recibe el nombre de *Network Triggered Service Request*.

En la Figura 3.12 se ilustra el procedimiento de petición de servicio originado por la red, cuyos pasos se describen a continuación.

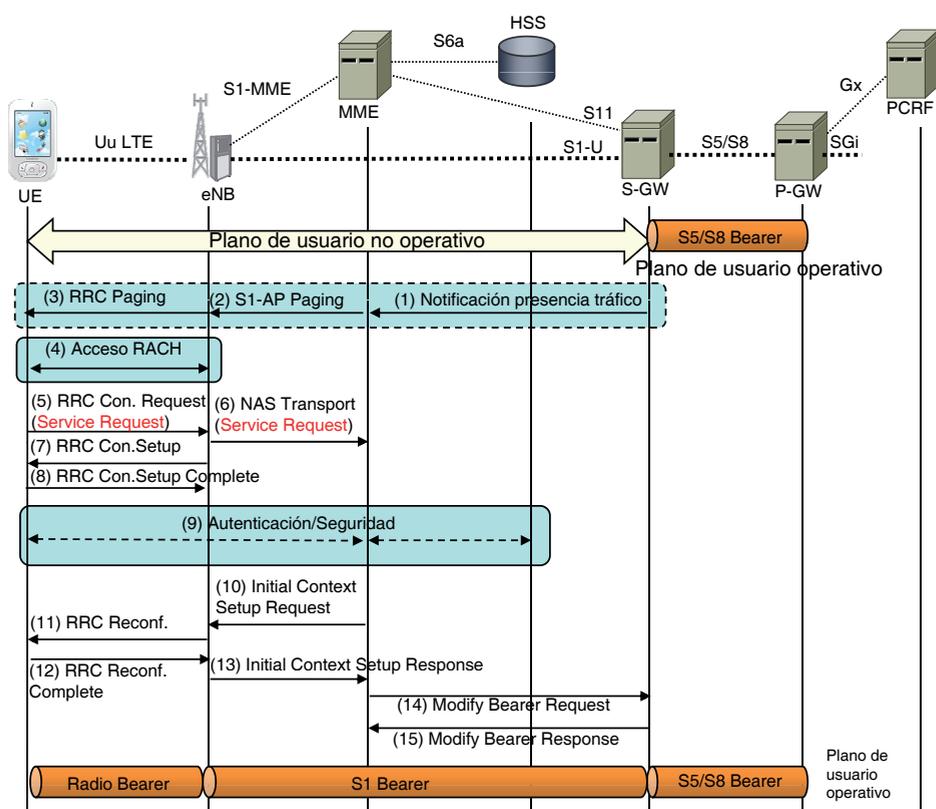


Figura 3.12 Procedimiento de petición de servicio

(1) - (3) Procedimiento de aviso (*Paging*)

En el ejemplo mostrado en la Figura 3.12, el procedimiento de petición de servicio lo inicia la red LTE cuando se reciben paquetes IP en la pasarela S-GW dirigidos a un usuario cuyo plano de usuario no se encuentra completamente operativo (si un usuario está en modo *idle*,

el plano de usuario solo está operativo entre las pasarelas P-GW y S-GW). En este caso, la pasarela S-GW notifica de la presencia de tráfico en el enlace descendente a la entidad MME que mantiene el registro del usuario. La entidad MME inicia entonces el procedimiento de aviso (*Paging*) mediante el envío del mensaje de aviso correspondiente mediante el protocolo S1-AP a todos los eNBs que pertenecen al área de seguimiento donde se encuentra registrado el usuario. Finalmente, el mensaje de *paging* es recibido por el terminal a través del canal PCCH (véase el Capítulo 5 para una descripción de los canales lógicos de la interfaz radio) de la celda a la que se encuentre sintonizado.

Los pasos (1), (2) y (3) no aplicarían en el caso de que fuera el terminal quien iniciara el procedimiento de petición de servicio debido a, por ejemplo, presencia de tráfico en el enlace ascendente.

(4) - (8) Establecimiento de la conexión RRC y envío del mensaje de petición de servicio a la red troncal EPC

El mensaje "*Service Request*" correspondiente a los protocolos NAS se envía encapsulado en mensajes RRC durante el procedimiento de establecimiento de una conexión RRC con el eNB. Al igual que se ha visto en el procedimiento de registro descrito en el apartado 3.2.5.1, el envío del mensaje "*Service Request*" comporta la utilización previa del procedimiento de acceso aleatorio y el establecimiento de la conexión RRC de forma que quedan activados los servicios portadores de señalización (i.e., SRB) a través de los cuales se soportan los posteriores intercambios de señalización NAS. El mensaje RRC que transporta el mensaje "*Service Request*" contiene también un identificador temporal del usuario (i.e., S-TMSI¹²) a partir del cual el eNB puede conocer la entidad MME que gestiona su acceso. Así, el mensaje "*Service Request*" es reenviado desde el eNB hacia la entidad MME correspondiente a través del protocolo S1-AP de la interfaz S1-MME. La recepción del mensaje "*Service Request*" en la entidad MME hace que el estado de movilidad del terminal conmute de estado *idle* a conectado (el modelo de estados de movilidad se describe en el apartado 3.3.1).

(9) Autenticación y seguridad

Los procedimientos de autenticación y activación de los servicios de seguridad podrían ejecutarse antes de proceder a la re-activación del plano de usuario.

(10) – (13) Establecimiento del contexto del usuario en el eNB y activación de los servicios portadores en la interfaz radio

La entidad MME envía el mensaje de control "*Initial Context Setup Request*" que contiene, entre otros, los parámetros de los servicios portadores EPS existentes (identificadores y parámetros de QoS), posibles restricciones de movilidad que tenga el usuario (i.e., lista de restricciones de *handover*), los parámetros de seguridad utilizados en la operativa de la interfaz radio y los parámetros necesarios para establecer el plano de usuario entre el eNB y el S-GW

¹² El identificador S-TMSI es un fragmento del identificador GUTI (que es más extenso) que permite identificar a un usuario (TMSI) en el contexto de una entidad MME y a la entidad MME correspondiente de entre las que dan servicio a través del eNB en cuestión (no de forma universal, como se hace en el GUTI).

(i.e., dirección IP de la pasarela S-GW e identificador de túnel TEID). A raíz de dicha petición, el eNB envía el mensaje “RRC *Connection Reconfiguration*” para re-establecer los servicios portadores radio. Los servicios de seguridad en el plano de usuario entre el equipo de usuario y el eNB también se re-establecen en este paso. Después del mensaje (12) el plano de usuario en la dirección ascendente ya se encuentra operativo, no así el descendente que todavía requiere el establecimiento del plano de usuario entre el eNB y el S-GW.

(14) y (15) Finalización de la configuración del servicio portador EPS

El plano de usuario en el enlace descendente se completa una vez el S-GW dispone de los parámetros que caracterizan la terminación del túnel GTP en el eNB. Dichos parámetros se envían en el mensaje de respuesta a la creación del contexto en el eNB (“*Initial Context Setup Response*”) y la entidad MME los señala a la pasarela S-GW mediante el procedimiento de modificación de las características de un servicio portador (mensaje “*Modify Bearer Request*”).

3.3. Gestión de movilidad

La gestión de movilidad es una de las piezas clave que caracteriza a los sistemas de comunicaciones móviles. Un requisito básico que deben satisfacer estos sistemas es permitir que los usuarios puedan acceder y recibir sus servicios desde cualquier ubicación geográfica donde el sistema disponga de cobertura, dejando aparte posibles limitaciones operativas o restricciones derivadas de las propias condiciones de uso de los servicios.

Por un lado, este requisito implica que el sistema de comunicaciones móviles tiene que albergar mecanismos que le permitan avisar a los usuarios de la activación de servicios originados desde la red dondequiera que se encuentren. Dada la gran extensión geográfica que puede abarcar una red celular, el envío de avisos (función de *paging*) a los terminales debe hacerse de forma “selectiva” a través únicamente de aquellas estaciones base donde exista una cierta probabilidad de encontrar al usuario. Para ello, el sistema debe hacer un “seguimiento” que le permita acotar la localización de los usuarios dentro de la zona de servicio de la red. Esta funcionalidad se conoce como gestión de la localización.

Por otro lado, cuando los usuarios se encuentran conectados al sistema a través de una determinada estación base, se requiere que el sistema sea capaz de mantener las conexiones activas aún cuando el terminal se encuentre en movimiento y resulte necesario realizar, en el transcurso de una conexión activa, un cambio de la estación base que le proporciona el acceso a la red. Esta funcionalidad se conoce como traspaso o *handover*.

Por tanto, la gestión de movilidad en un sistema de comunicaciones móviles abarca tanto la gestión de la localización como la gestión del *handover*. Asimismo, las funciones propias de gestión de movilidad deben complementarse con funciones que permitan la autenticación de los usuarios y la autorización de acceso a los servicios solicitados desde cualquier estación base a través de la que se conecte el terminal. Las funciones de autenticación y au-

torización se abordan en el apartado 3.4.2 en el contexto del marco de seguridad de acceso a la red LTE.

En los siguientes apartados se describe el marco de gestión de movilidad considerado en un sistema LTE.

3.3.1 Marco de gestión de movilidad

El marco de gestión de movilidad desarrollado en el sistema LTE cubre los siguientes escenarios:

- Movilidad intra-LTE. Gestión de movilidad específica para el acceso a través de E-UTRAN.
- Movilidad entre redes de acceso 3GPP, es decir entre E-UTRAN, UTRAN y GERAN. El sistema LTE incorpora mecanismos específicos para gestionar la localización así como la realización de *handovers* (denominados *inter-RAT handovers*) entre las diferentes redes de acceso 3GPP. Las soluciones de *interworking* y la gestión de movilidad asociada se abordan en el apartado 7.3 del Capítulo 7 para escenarios de despliegue donde co-existan redes de acceso UTRAN y/o GERAN con E-UTRAN.
- Movilidad con redes de acceso no especificadas por 3GPP, como por ejemplo redes CDMA2000, WLAN o Mobile WiMAX. En este caso, el sistema LTE incluye soporte para garantizar la continuidad de servicio entre estas redes mediante un mecanismo de *handover* entre sistemas. Las soluciones de *interworking* entre redes 3GPP y no 3GPP y la gestión de movilidad asociada se aborda en el apartado 7.4 del Capítulo 7.

El marco de gestión de movilidad cubierto en este apartado se centra básicamente en el escenario de movilidad intra-LTE mientras que, tal como se ha mencionado previamente, la movilidad entre E-UTRAN y otras redes de acceso se cubre en el Capítulo 7.

El sistema LTE especifica un modelo de movilidad (denominado modelo EMM, *EPS Mobility Management*) con dos posibles estados que representan dos situaciones de accesibilidad en las que puede encontrarse un usuario del sistema. Los dos estados del modelo EMM son [4]:

- Estado "No registrado" (*EMM-Deregistered*). En este estado, el usuario no se encuentra visible en el sistema LTE, y por tanto, no tiene acceso a los servicios del sistema. En este estado, el sistema LTE no dispone de ninguna información relativa a la localización del usuario. La situación más habitual que conlleva que un usuario se encuentre en este estado es que el terminal LTE está apagado.
- Estado "Registrado" (*EMM-Registered*). En este estado, el usuario está operativo en el sistema LTE, y por tanto, tiene acceso a sus servicios a través de un equipo de usuario. En este estado, la red dispone de información de localización del equipo de usuario con la resolución de, al menos, una lista de áreas de seguimiento (*Tracking Area list*). Asimismo, en este estado el equipo de usuario mantiene, al menos, una conexión PDN

activa y tiene asignado un nodo MME de la red troncal encargado de realizar su seguimiento.

La conmutación entre ambos estados de movilidad se realiza a través de procedimientos de "registro" y "cancelación de registro". El procedimiento de registro, descrito en detalle en el apartado 3.2.5.1 en el contexto de procedimientos de gestión de sesiones, conlleva el paso de un estado "No registrado" a un estado "Registrado" y siempre es un procedimiento iniciado por el terminal. El procedimiento de "cancelación de registro" (i.e., *Network Detach*) sirve para realizar el paso contrario y puede ser iniciado tanto por el terminal (e.g., en el proceso de apagado del terminal) como por la propia red (e.g., cambios en la suscripción de un usuario). La conmutación de un estado "Registrado" a "No registrado" también puede acontecer debido a otros motivos tales como el rechazo de la red al registro de un terminal en una determinada área de localización o bien el hecho de que, pasado un cierto tiempo, no se haya recibido ninguna actualización por parte del terminal (la red puede forzar a que los equipos realicen actualizaciones periódicas de su localización aún cuando no cambien de área de seguimiento).

Junto con los estados de movilidad EMM, el sistema LTE también define un modelo de estados para indicar la existencia o no de un plano de control activo entre el equipo de usuario y el nodo MME de la red troncal donde se encuentra registrado. Dicho modelo de estados se denomina modelo ECM (*EPS Connection Management*) y se estructura también en dos posibles estados:

- Estado "Desconectado" (*ECM-Idle*). En este estado, el terminal no tiene establecida una conexión de señalización con ninguna entidad MME. La existencia de este estado responde básicamente a la necesidad de disponer de un modo de operación de bajo consumo que permita conseguir un modelo de funcionamiento "*Always on*" mediante la posibilidad de conmutar de forma rápida entre este estado y el siguiente estado "Conectado" en el que el terminal podría enviar/recibir datos.
- Estado "Conectado" (*ECM-Connected*). En este estado, el equipo de usuario tiene establecida una conexión de señalización con una entidad MME de la EPC. Dicha conexión de señalización se compone de una conexión RRC en E-UTRAN y de una conexión a través de la interfaz S1-MME entre la red de acceso E-UTRAN y la entidad de la red troncal MME. El envío/recepción de datos de usuario siempre se realiza en este estado.

La información relativa a los estados EMM y ECM de un usuario se almacena en el equipo de usuario y en la red troncal EPC, en particular en el nodo MME que da servicio a dicho equipo terminal. En el caso del terminal, los estados *ECM-Idle* y *ECM-Connected* se mapean directamente a los estados *RRC_IDLE* y *RRC_CONNECTED* empleados por la capa RRC del plano de control entre el terminal y E-UTRAN (en el apartado 5.2 del Capítulo 5 se describen los estados RRC). En el caso del nodo MME, el estado *ECM-Connected* se vincula a la existencia de una conexión de señalización en la interfaz S1-MME asociada al usuario entre el MME y el eNB correspondiente.

En la Figura 3.13 se ilustra el modelo de estados comentado, junto con el conocimiento

que la red troncal EPC tiene en cada estado respecto a la localización del terminal y los mecanismos utilizados para controlar la movilidad.

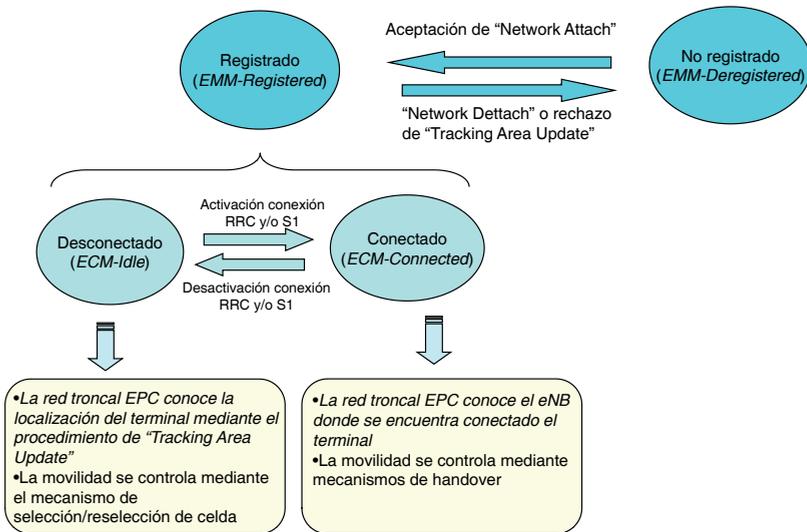


Figura 3.13 Modelo de estados para la gestión de movilidad y de la conexión a EPC

Cuando un equipo de usuario se encuentra en estado "Registrado" y "Desconectado", el equipo de usuario va decodificando la información recibida a través de los canales de control de las estaciones base del sistema y ejecuta los mecanismos de selección/re-selección de celda [6] y de selección de red [7]. En este estado, la red de acceso E-UTRAN no dispone de ninguna información relativa al equipo de usuario (no hay ningún contexto activo con información del usuario en E-UTRAN) y el seguimiento de su localización dentro de la zona de servicio del sistema se realiza desde la red troncal EPC a través de un procedimiento de actualización del área de seguimiento (*Tracking Area Update*, TAU). La gestión de la localización en LTE mediante el procedimiento TAU se explica más detalladamente en el apartado 3.3.2, mientras que las principales características del mecanismo de selección/re-selección que controla la movilidad entre celdas se proporcionan en el apartado 6.6.2 del Capítulo 6 dedicado a la gestión de recursos radio. Encontrándose en el estado "Registrado" y "Desconectado", el equipo de usuario debe responder a los mensajes de aviso (*paging*) enviados desde un nodo MME a través de las estaciones base que integran el área de seguimiento donde está registrado el equipo. La respuesta a un aviso se realiza a través del envío de un mensaje de petición de servicio (*Service Request*), que es el mismo mensaje utilizado en el caso de que el equipo de usuario solicite a la red el establecimiento de los servicios portadores para el envío de datos en el canal ascendente (el procedimiento de petición de servicio se detalla en el apartado 3.2.5.3). En particular, entre los requerimientos descritos en el informe técnico TR 25.913 [8] elaborado por el 3GPP se indica que la latencia para pasar de un modo "Desconectado" a un modo "Conectado", y por tanto pasar de no tener recursos asignados en E-UTRAN

a disponer del plano de usuario establecido y preparado para enviar/recibir información, debe ser inferior a 100 ms.

Cuando un equipo de usuario se encuentra en estado “Registrado” y “Conectado”, el nodo MME conoce su localización a nivel de cuál es el eNB que le está dando servicio en E-UTRAN (i.e., el eNB con el que tiene la conexión RRC). En esta situación, la movilidad del equipo de usuario se controla mediante el procedimiento de *handover*. Más detalles sobre los diferentes tipos de mecanismos de *handover* soportados en LTE se proporcionan en el apartado 3.3.3. Asimismo, en este estado, el equipo de usuario sigue ejecutando el procedimiento de actualización del área de seguimiento (i.e., TAU) en caso de que sea necesario.

3.3.2 Gestión de la localización

Tal como se ha descrito en el apartado anterior, la red debe disponer de mecanismos para conocer con un determinado nivel de resolución la localización de los terminales que se encuentren registrados (estado EMM-*Registered*) pero que no tengan establecida una conexión con ninguna estación base (estado ECM-*Idle*).

Para ello, en LTE se define el concepto de Área de Seguimiento (*Tracking Area*, TA) para gestionar la información de localización. Un TA agrupa a un conjunto de eNBs de forma que la información de localización disponible en la red troncal EPC de un determinado equipo de usuario solamente se conoce en base a la resolución proporcionada por tales agrupaciones. La identidad de un TA se denomina TAI (*Tracking Area Identifier*) y se difunde a través de los mensajes de información de sistema enviados en los canales de *broadcast* de los eNBs que integran una TA. Un eNB sólo puede pertenecer a una TA de una red troncal EPC, es decir, no hay solapes entre diferentes TAs. El equipo de usuario, a partir del identificador TAI recibido, es el encargado de comunicar a la red en qué TA se encuentra accesible mediante los mecanismos de *Network Attach* y de *Tracking Area Update*. De esta forma, cuando la red necesita contactar con el terminal, el mensaje de aviso (*paging*) se difunde a través de todas las estaciones base que integran el TA en que se encuentra localizado el terminal.

La selección de qué estaciones base integran cada una de las TAs es un aspecto de ingeniería que plantea un compromiso entre la carga de señalización que originan los mecanismos de TAU y la carga asociada al procedimiento de aviso de los equipos de usuario: si las TAs están formadas por pocos eNBs, el número de accesos a la red por parte de los terminales para actualizar su localización aumenta. Un elevado número de accesos repercute tanto en la carga de la red como en el consumo de baterías de los equipos de usuario. En cambio, cuando la red debe avisar a un equipo de usuario, el procedimiento de *paging* puede ser más selectivo ya que afecta a pocas celdas.

La gestión de la localización mediante áreas de seguimiento es un planteamiento ya utilizado en redes anteriores como UMTS y GSM. En particular, en dichas redes, la localización de terminales registrados en el dominio de circuitos se realiza mediante el concepto de área de localización (Location Area, LA) y mediante la definición de áreas de rutado (*Routing Area*,

RA) en el dominio de paquetes¹³. Al igual que las TA de LTE, las agrupaciones LA y RA tampoco presentan solape y su determinación plantea los mismos compromisos esgrimidos anteriormente. Sin embargo, una modificación importante que introduce LTE es la posibilidad de que un terminal se encuentre registrado en múltiples TAs a la vez, es decir, en una lista de TAs. Esta modificación responde a la necesidad de disponer de mecanismos que ayuden a minimizar la señalización debida a las actualizaciones de localización en las zonas limítrofes entre TAs. En estas zonas, determinados patrones de movilidad de los equipos de usuario, podrían originar un número de TAUs elevado si el terminal va conmutando sucesivamente entre eNBs de diferentes TAs. Para ello, mediante el registro a múltiples TAs, el equipo terminal únicamente actualiza su localización en el caso de que se mueva fuera de la cobertura del grupo de TAs que integran la lista, con la excepción de la actualización periódica de TA. El conjunto de TAs que integran la lista lo controla la red y lo comunica al terminal mediante el procedimiento de actualización de TA (cuando el terminal se registra en una determinada TA, la red le envía la lista de TAs en las que mantiene un registro para el terminal). Nótese que, mediante una gestión apropiada de la lista de TAs por terminal, la red LTE puede confeccionar de forma dinámica áreas de localización “personalizadas” que conduzcan a una reducción de la señalización asociada a los procedimientos de TAU y *Paging*.

Es importante destacar que el marco de gestión de movilidad en LTE ha sufrido importantes simplificaciones respecto a la solución de movilidad de UMTS donde la red de acceso UTRAN también participa en la gestión de la localización mediante la definición de áreas de rutado propias (UTRAN Routing Areas, URAs) y de un conjunto de estados asociados a la gestión de movilidad en la capa RRC (e.g., cell-FACH, URA-PCH, etc.). En cambio, en el sistema LTE, la gestión de localización se sustenta íntegramente en la red troncal EPC.

3.3.3 Mecanismo de handover

El mecanismo de *handover* se utiliza para gestionar la movilidad de los equipos de usuario que se encuentran en modo activo (*ECM-Connected*). Conceptualmente, el *handover* es un mecanismo que permite que las conexiones que tengan establecidas los equipos de usuario “sobrevivan” al cambio de estación base que proporciona el acceso a la red.

Desde la perspectiva del servicio ofrecido al usuario, los requisitos de diseño de un mecanismo de preparación y ejecución del *handover* se plantean en términos del tiempo de interrupción o tasa de pérdida de datos que puede aparecer durante la ejecución del cambio. En este sentido, en los requerimientos de diseño del sistema LTE [8], se establece que la degradación de prestaciones en la que puede incurrirse durante la realización de un *handover*

¹³ El motivo de que existan dos áreas diferentes de seguimiento en UMTS (i.e., LA y RA) en función del dominio de la red troncal al que se acceda obedece al hecho de que el compromiso entre señalización de *paging* y señalización de actualizaciones presenta un punto de operación óptimo diferente entre el tráfico de circuitos y el tráfico de paquetes. Así, en los servicios ofrecidos a través del dominio de circuitos, el envío de un aviso se realiza únicamente al inicio del servicio. Por el contrario, en los servicios de paquetes, el canal de *paging* sirve también para señalar la presencia de tráfico de bajada en la red pendiente de ser enviado en los casos en los que el terminal se encuentre en modo *idle*. En cualquier caso, en UMTS un área de rutado RA siempre se encuentra contenida en un área de localización LA, la cual puede albergar una o múltiples áreas de rutado RA.

debe ser menor o igual a la existente en redes de circuitos GSM. Asimismo, entre los requerimientos de velocidades físicas de los terminales, se apuntan velocidades de hasta 350 km/h, aunque se indica que la red esté realmente optimizada para trabajar en el rango 0-15 km/h. Todas estos requerimientos hacen que el diseño del mecanismo de *handover* en LTE constituya un elemento clave del sistema.

Para la consecución de tales requisitos, la implementación del mecanismo de *handover* en LTE se sustenta en los siguientes principios básicos:

- El mecanismo de *handover* especificado en LTE se controla desde la red (la red decide el cambio de estación base) teniendo en cuenta medidas enviadas desde los equipos de usuario (*handover* controlado por la red y asistido por el terminal). Concretamente, la decisión de llevar a cabo un cambio de eNB de un terminal en modo conectado, la toma el propio eNB con el que el equipo de usuario mantiene una conexión RRC activa. En el apartado 6.6 del Capítulo 6 se proporcionan más detalles del control del mecanismo de *handover* (criterios y algoritmos de decisión), atendiendo a que la función que controla las decisiones de *handover* constituye una de las funciones clave para la gestión eficiente de los recursos radio en E-UTRAN.
- Una vez tomada la decisión sobre la necesidad de realizar un *handover*, la propia red se encarga de reservar los recursos necesarios en el eNB destino en aras a reducir el tiempo de interrupción (durante este tiempo de preparación el terminal sigue conectado al eNB antiguo) así como el número de intentos de *handover* fallidos. Una vez la red garantiza que el terminal puede ser traspasado al nuevo eNB, la red ordena al terminal que ejecute el cambio. Este planteamiento se conoce popularmente como *Make Before Break*, en contraposición a otro planteamiento denominado *Break Before Make* donde el terminal realizaría el cambio de eNB sin haberse efectuado ninguna reserva de recursos en el eNB destino.
- Durante la realización del proceso de *handover*, la propia red dispone de mecanismos para transferir los paquetes del usuario pendientes de transmisión en el viejo eNB hacia el eNB destino (e.g., mediante la utilización de la interfaz X2). Este planteamiento permite reducir el número de paquetes perdidos durante la ejecución de un *handover*. Nótese por ejemplo que, a diferencia de la red de acceso UTRAN basada en CDMA, el *handover* en E-UTRAN siempre es del tipo "*hard-handover*" (en contraposición al denominado "*soft-handover*" característico de sistemas CDMA). Un "*hard-handover*" significa que existe un tiempo de interrupción del servicio durante el cambio de estación base necesario para adquirir la sincronización en la nueva base y obtener la primera asignación de recursos en ella (en sistemas CDMA, no se produce dicho tiempo de interrupción). Además, tal como se verá más detalladamente en el apartado de procedimientos de movilidad, en el momento en que un terminal conmuta entre estaciones base, todavía pueden seguir llegando paquetes IP a la estación base antigua dado que el rutado en la infraestructura de la red todavía puede no haberse actualizado.

Sobre estas premisas, la realización de un handover intra-LTE puede dar respuesta a diferentes escenarios de movilidad tales como:

- *Handover* entre eNBs conectados mediante una interfaz X2. La existencia de esta interfaz, tal como se describe en el apartado 2.5.2 del Capítulo 2, permite establecer un plano de usuario entre eNBs para el envío de datos durante el proceso de *handover*. Además, a través de X2, la señalización del procedimiento así como la transferencia del contexto de datos asociado al equipo terminal puede llevarse a cabo directamente entre eNBs, sin pasar por el nodo MME de la red troncal.
- *Handover* entre eNBs que no disponen de la interfaz X2. En este caso, no es posible el envío de paquetes de usuario entre eNBs y la señalización de *handover* debe articularse necesariamente a través de la entidad MME.
- *Handover* entre eNBs, soporten o no soporten la interfaz X2, que requiera la reubicación de alguno de los nodos de la troncal EPC. En este caso, el cambio de eNB podría conllevar el cambio de la pasarela S-GW a través de la cual está establecido el plano de usuario o bien del nodo MME que termina el plano de control con el equipo de usuario. El caso más complejo sería el de un *handover* donde se cambiaran ambos, S-GW y MME.

Es importante destacar que el mecanismo de *handover* no considera en ningún caso el cambio de la pasarela P-GW que se mantiene como el punto de conexión a la red externa durante toda la vigencia de la conexión PDN.

3.3.4 Procedimientos de gestión de movilidad

Los principales procedimientos asociados a la gestión de movilidad soportados en el sistema LTE son los siguientes:

- Procedimiento de registro (*Network Attach*). Tal como se ha comentado en el apartado 3.2.5, el procedimiento de registro es un procedimiento combinado de gestión de movilidad y gestión de sesiones. Sus detalles se analizan en el apartado 3.2.5.1.
- Procedimiento de actualización de áreas de seguimiento (*Tracking Area Update*). El procedimiento tiene diferentes variantes en función de si la actualización conlleva el cambio de las pasarelas MME y/o S-GW de la red troncal.
- Procedimientos de *handover*. El sistema LTE contempla diferentes procedimientos de *handover* en E-UTRAN en función del soporte de la interfaz X2 y de la necesidad de cambiar las pasarelas S-GW y MME durante el proceso. Asimismo, se definen procedimientos de handover específicos entre E-UTRAN y otras redes de acceso. Más detalles sobre el mecanismo de *handover* entre diferentes tipos de tecnologías se proporciona en el Capítulo 7.

A continuación se describe el procedimiento de actualización de áreas de seguimiento y

los procedimientos de *handover* en el caso de utilizarse la interfaz X2 y en el caso de que ésta no exista y el *handover* se sustente íntegramente desde la red troncal EPC a través de la interfaz S1. La descripción de estos tres procedimientos básicos proporciona una visión bastante completa del alcance de la gestión de movilidad en el sistema LTE. El lector interesado en profundizar más en la operativa de estos procedimientos, o bien en el resto de procedimientos de gestión de movilidad no abordados de forma explícita en el libro, puede encontrar su descripción en la especificación 3GPP TS 23.401 [4].

3.3.4.1 Actualización del área de seguimiento

En la Figura 3.14 se ilustra el procedimiento de actualización del área de seguimiento. El caso mostrado es el más complejo ya que considera que la actualización de la localización conlleva tanto un cambio de la entidad MME que termina la señalización NAS con el usuario móvil como de la pasarela S-GW por donde transcurre el plano de usuario en la red troncal EPC. Los casos más simples donde se mantiene la entidad MME y/o la pasarela P-GW siguen el mismo procedimiento exceptuado algunos pasos intermedios que se indican durante la descripción del procedimiento completo.

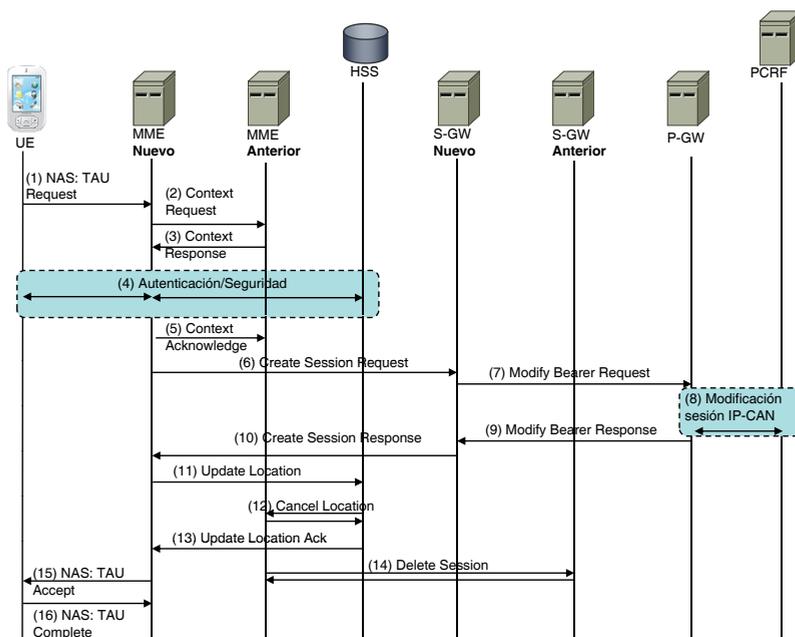


Figura 3.14 Procedimiento de actualización del área de seguimiento

Los pasos mostrados en la Figura 3.14 se describen a continuación:

(1) Petición de actualización de TA

El procedimiento de actualización de TA lo inicia el equipo de usuario cuando, por ejem-

plo, sintoniza una nueva celda que señala a través de los canales de control radio (i.e., canales BCCH) un identificador de TA no incluido en su lista de TAs registradas.

El terminal inicia la actualización de TA mediante el envío de un mensaje NAS “*TAU Request*” a través del eNB al que se encuentra sintonizado. El mensaje NAS contiene, entre otros parámetros, un identificador del usuario (e.g., GUTI asignado previamente en el proceso de registro o en una actualización de TA anterior). En la interfaz radio, el mensaje RRC en el que se transporta encapsulado el mensaje NAS indica al eNB la identidad de la entidad MME en la que el terminal se encuentra actualmente registrado (e.g., identificador GUMMEI). Nótese que, en el caso representado en la Figura 3.14, sucedería que la entidad MME anterior no estaría asociada al nuevo eNB, por lo que éste encaminaría el mensaje NAS hacia una MME nueva que formara parte de su *pool area*¹⁴.

(2) – (5) Transferencia de contextos entre MMEs

Tras la recepción del mensaje NAS, la nueva entidad MME utiliza el identificador de usuario GUTI para extraer la identidad de la MME anterior en la que el usuario se encontraba registrado. De esta forma, la MME nueva contacta con la MME anterior para recuperar el contexto de información asociado a dicho usuario. En el proceso de obtención del contexto, la MME anterior verifica que la petición de TAU recibida en la nueva MME sea válida y, en caso afirmativo, se produce la transferencia del contexto.

En caso de que hubiera algún problema con la verificación y no pudiera llevarse a cabo la recuperación del contexto anterior, la entidad MME nueva iniciaría los procedimientos de autenticación y seguridad descritos posteriormente en el apartado 3.4.4. Nótese que estos mecanismos son siempre necesarios en los casos donde en la entidad MME que cursa la actualización no exista un contexto del usuario previamente establecido (o recuperado de la entidad MME previa en la que el usuario en cuestión estuviera registrado) con una asociación de seguridad válida.

El mecanismo de transferencia de contextos y los mecanismos de seguridad no serían necesarios en el caso de que la actualización de TA no comportara el cambio de MME.

(6) – (10) Cambio de S-GW y modificación del plano de usuario entre pasarelas

Estos pasos son necesarios únicamente cuando la actualización de TA comporta además un cambio de la pasarela S-GW donde se ancla el plano de usuario en la red troncal EPC. Para ello, la entidad MME inicia el restablecimiento de los servicios portadores EPS que el terminal tuviera activos a través de la pasarela S-GW nueva. En el mensaje “*Create Session Request*” se proporciona la información necesaria (extraída del contexto del usuario recuperado del MME anterior) para replicar el estado de los servicios portadores que se encuentran (todavía) configurados en la pasarela S-GW anterior. Entre la información proporcionada se encuentra la identidad de la pasarela P-GW que proporciona el servicio de conectividad. El S-GW nuevo utiliza esta información para interactuar con la pasarela P-GW (mensajes (7) y (8)) y establecer el plano de usuario entre ellos.

¹⁴ El pool area de un eNB son todas las entidades MME a las que tiene acceso.

Opcionalmente, en el paso (8), la pasarela P-GW podría también informar a la entidad PCRF del subsistema PCC del cambio de localización (y si fuera el caso, el cambio de red de acceso). Esta información podría originar el cambio de reglas PCC en casos donde, por ejemplo, se aplicaran mecanismos de tarificación en función de la localización de los usuarios.

(11) – (13) Actualización de la localización en el HSS

La entidad MME nueva informa a la base de datos HSS de la nueva localización del usuario. Posteriormente, la base de datos HSS contacta con el nodo MME anterior para proceder a la eliminación del contexto del usuario en dicho nodo. La eliminación del contexto puede retardarse un cierto tiempo mediante la programación de un contador. Esto permite que, en caso fallido del procedimiento, o bien en caso de que el terminal proceda a realizar una nueva actualización de TA hacia la MME anterior de forma rápida, la MME anterior todavía disponga del contexto y se evita la transferencia del mismo.

(14) Eliminación de contextos en la pasarela S-GW antigua

La entidad MME anterior, en paralelo al proceso de eliminación del contexto del usuario comentado en el paso previo, es la encargada iniciar la eliminación el contexto del terminal existente la pasarela S-GW anterior con la configuración de los servicios portadores EPS.

(15) y (16) Finalización del procedimiento de actualización de TA

El procedimiento de actualización de TA termina con el envío del mensaje NAS "TAU Accept" al terminal donde se le indica su nuevo identificador GUTI y la lista de TAs en la que se encuentra registrado. Finalmente el terminal confirma el procedimiento enviado en el mensaje "TAU complete".

3.3.4.2 Handover con soporte de la interfaz X2

Este procedimiento es utilizado para transferir la conexión de un equipo de usuario entre dos eNBs que se encuentren conectados entre ellos mediante una interfaz X2. El procedimiento se ilustra en la Figura 3.15 para el caso en que las entidades de la red troncal MME y S-GW no cambian en el proceso. El punto de partida del procedimiento representado es el de un terminal UE que mantiene una conexión RRC (i.e., el terminal se encuentra en modo RRC_CONNECTED, y por extensión, en modo *ECM-Connected*) con el eNB origen y tiene activado uno o varios servicios portadores EPS con la red LTE. A continuación se describen los principales pasos del procedimiento.

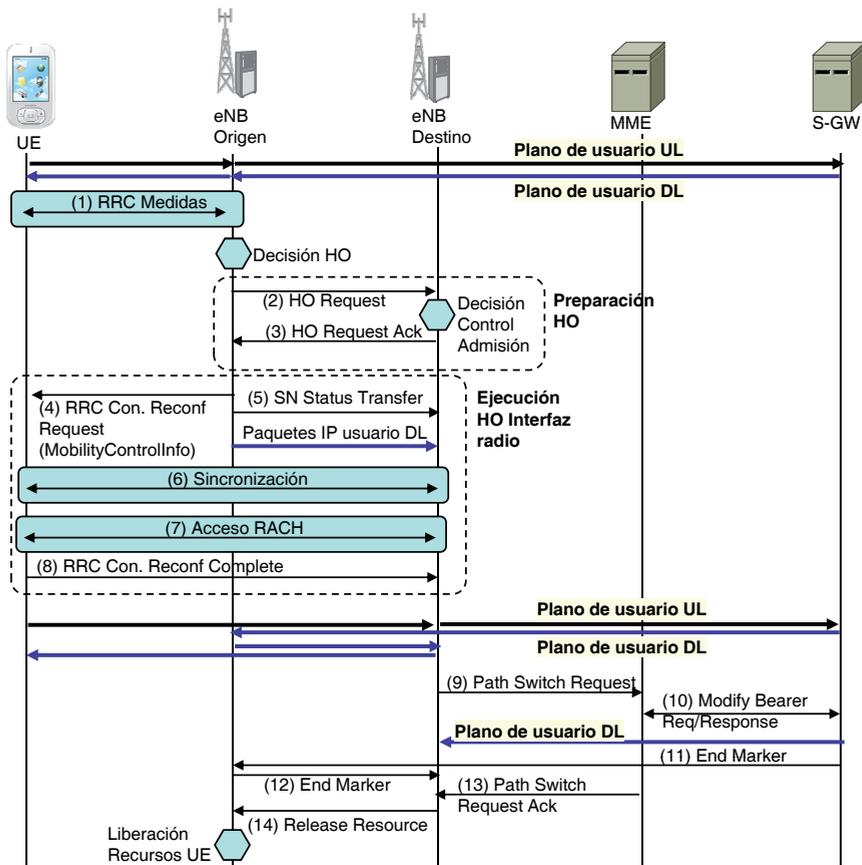


Figura 3.15 Procedimiento de handover con soporte de la interfaz X2 (X2-based handover)

(1) Iniciación del *handover*

En E-UTRAN el procedimiento de *handover* (HO) siempre lo inicia el eNB con el que se encuentre conectado el terminal. Para asistir en la toma de decisiones, el eNB puede ordenar que el equipo de usuario realice ciertas medidas (e.g., niveles de potencia recibidos desde eNBs vecinos) y las envíe a través de la conexión RRC al eNB. El algoritmo utilizado para tomar la decisión de HO no forma parte de las especificaciones 3GPP y puede tomar en consideración múltiples aspectos (véase el apartado 6.6 del Capítulo 6 para una descripción más detallada del proceso y criterios de decisión de *handover*).

(2) y (3) Fase de preparación de HO

Una vez el eNB origen ha decidido que la conexión debe traspasarse al eNB destino, el eNB origen envía la petición de HO ("*Handover Request*") al eNB destino a través de la interfaz X2. La petición de HO contiene una descripción de los servicios portadores radio que el terminal tiene activados en el eNB origen, junto con el resto de información almacenada en el contex-

to de datos que el usuario tiene en el eNB (e.g., identificador radio, parámetros de seguridad, configuración de los portadores de señalización con la MME que gestiona el terminal, etc.). Con esta información, el eNB destino puede realizar el control de admisión y reservar los recursos de transmisión necesarios (e.g., para servicios portadores con tasa garantizada). En el caso ilustrado en la Figura 3.15, el eNB destino indica que puede acomodar la conexión mediante el mensaje "*Handover Request Ack*". El mensaje de respuesta puede contener un nuevo identificador radio (i.e., C-RNTI) y un preámbulo RACH reservados para el usuario en el eNB destino.

(4) – (8) Ejecución del HO en la interfaz radio

La ejecución del HO empieza con el envío de un mensaje RRC (paso (4)) al equipo de usuario donde se ordena el cambio de eNB y se transfiere la información pertinente proporcionada por el eNB nuevo (campo *mobilityControlInformation*). Al mismo tiempo, el eNB origen puede enviar información (mensaje (5)) relativa al estado del plano de usuario de los servicios portadores en el momento del cambio (e.g., números de secuencia del protocolo PDCP) en aras a preservar la ordenación del tráfico en los servicios que lo requieran. A partir de este momento, los paquetes IP pendientes de transmisión en el eNB origen, o que sigan recibándose en éste, se reenvían al eNB destino a través de la interfaz X2.

En la interfaz radio, el terminal sincroniza los canales de control de la celda destino y accede a ella mediante el canal de acceso aleatorio (RACH). En caso de que el terminal haya recibido un preámbulo RACH dedicado, el acceso al canal RACH se realiza sin contención (en caso contrario, el terminal accede al RACH siguiendo el mecanismo basado en contención). Como parte del mecanismo de acceso al canal RACH, el terminal recibe el ajuste del avance temporal (*Time Advance*) en el eNB destino y una asignación inicial de recursos radio.

Finalmente, mediante el mensaje (8), el terminal señala al eNB destino la realización del cambio identificándose mediante el nuevo C-RNTI. A partir de este instante los servicios portadores de la interfaz radio vuelven a estar operativos para la transferencia de información (tanto de control como datos de usuario). Según el documento 3GPP TS 36.133 [9], el tiempo de interrupción máximo que puede transcurrir desde la recepción del mensaje (4) hasta el acceso en el canal RACH del eNB destino (paso (7)) se encuentra entre los 50 y 130 ms¹⁵. A este tiempo hay que sumarle el tiempo de acceso al canal RACH y el envío del primer mensaje de señalización (8) que puede representar unas decenas de milisegundos adicionales.

(9) – (13) Actualización de las conexiones con la red troncal EPC

Una vez el terminal se encuentra conectado a través del eNB nuevo, el plano de usuario en la red troncal todavía sigue establecido a través del eNB origen, tal como se ilustra en la Figura 3.15.

La actualización de las conexiones en la red troncal EPC se inicia mediante el envío del mensaje "*Path Switch Request*" desde el eNB destino hacia la entidad MME correspondien-

¹⁵ El tiempo de interrupción depende en gran medida del tiempo de búsqueda de la celda destino.

te. El mensaje incluye la identificación de la nueva celda y, si fuera el caso, la lista de servicios portadores restablecidos en el nuevo eNB. En este punto la entidad MME decide si la pasarela S-GW puede mantenerse (como sería en el caso ilustrado en la Figura 3.15) o bien debe cambiarse.

En caso de no ser necesario el cambio de pasarela S-GW, la entidad MME interactúa con ella para actualizar el contexto de datos asociado al usuario y conmutar el túnel GTP en la interfaz S1 desde el eNB origen al eNB destino. Tan pronto como la pasarela procesa el mensaje (10), ésta empieza a enviar los paquetes IP del usuario hacia el eNB nuevo. En este punto, la pasarela S-GW puede enviar uno o varios paquetes denominados "*End Marker*" (son paquetes GTP-U que no transportan ningún paquete IP del usuario) hacia el eNB origen para que sepa que la conmutación de caminos ya se ha llevado a cabo. El eNB origen reenvía el mensaje "*End Marker*" al eNB destino a través del túnel GTP establecido en la interfaz X2. Tras la recepción de este paquete, el eNB destino sabe que el camino viejo ya está vacío y pueden eliminarse los recursos de transmisión entre eNBs.

Las acciones relacionadas con el proceso de HO en la red troncal finalizan tras el envío del mensaje de respuesta a la petición de "*Path Switch Request*".

(14) Liberación de recursos

Finalmente, el eNB destino notifica al eNB origen que el HO se ha realizado con éxito. A partir de este instante, el eNB origen puede eliminar el contexto de datos del usuario y, si no lo hubiera hecho ya, desactivar el re-envío de paquetes IP a través de la interfaz X2.

3.3.4.3 Handover sin soporte de la interfaz X2

Este procedimiento de *handover*, denominado como *handover* basado en S1, se utiliza en caso de que las estaciones base origen y destino no estén conectadas mediante una interfaz X2. La Figura 3.16 muestra la operativa de este procedimiento en el mismo escenario planteado para el análisis del *handover* basado en X2. En aras a facilitar la comparación de ambos procedimientos, los mensajes nuevos necesarios en el *handover* basado en S1 se resaltan en color rojo en la Figura 3.16. A continuación se describen los pasos de este procedimiento.

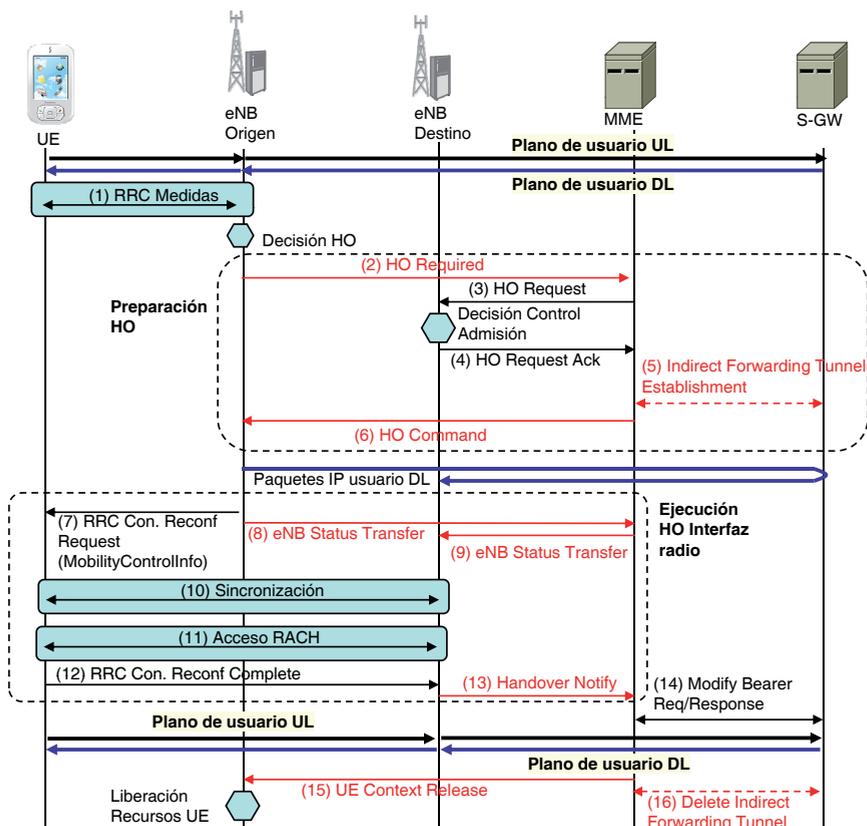


Figura 3.16 Procedimiento de handover sin soporte de la interfaz X2 (S1-based handover)

(1) Iniciación del handover

La iniciación del *handover* es idéntica a la descrita para el caso de handover basado en X2.

(2) - (6) Fase de preparación de HO

La diferencia fundamental respecto al *handover* basado en X2 es que, en este caso, la petición de *handover* la envía el eNB origen hacia la MME que termina el plano de control del terminal en cuestión a través de la interfaz S1. El mensaje contiene la información necesaria para que la entidad MME puede enviar la petición de *handover* (mensaje (3)) al eNB destino (nótese que en el HO basado en X2, el mensaje (3) se envía directamente entre los eNBs). Tras la recepción de la respuesta del eNB destino, la entidad MME ordena la ejecución del HO al eNB origen (mensaje (6)) y, opcionalmente, puede ordenar el establecimiento de un túnel GTP en la pasarela S-GW que permita la transferencia de paquetes de usuario pendientes de transmisión en el eNB origen hacia el eNB destino. Este mecanismo se denomina *Indirect*

Forwarding Tunnel y también puede utilizarse para el establecimiento de un túnel entre la pasarela S-GW origen y destino en el caso de que el procedimiento de *handover* comporte su cambio.

(7) –(13) Ejecución del HO en la interfaz radio

Esta fase es prácticamente igual que la fase de ejecución del HO descrita para el *handover* basado en X2. Las dos únicas diferencias son que el mensaje equivalente al mensaje “*SN Status Transfer*” representado en la Figura 3.15 se envía ahora entre eNBs a través de la entidad MME, y que, una vez detectado el acceso del terminal en el nuevo eNB, éste lo señala a la entidad MME mediante el mensaje “*Handover Notify*”.

(14) Actualización de las conexiones con la red troncal EPC

Recibida la notificación de HO desde el eNB destino, la entidad MME interactúa con la pasarela S-GW para actualizar el contexto de datos asociado al usuario y conmutar el túnel GTP en la interfaz S1 desde el eNB origen al eNB destino. Tan pronto como la pasarela procesa el mensaje (14), ésta empieza a enviar los paquetes IP del usuario hacia el eNB nuevo.

(15) y (16) Liberación de recursos

Finalmente, la entidad MME indica al eNB origen que el HO se ha realizado con éxito y que puede proceder a la eliminación del contexto y recursos asociados al usuario. También, en caso de haberse establecido en túnel GTP para la transferencia de datos entre el eNB origen y el eNB destino a través de la pasarela S-GW (mecanismo *Indirect Forwarding*), la entidad MME procede a su eliminación.

3.4 Gestión de seguridad

Los mecanismos para garantizar la seguridad de la información tienen un papel muy relevante en los sistemas de comunicaciones móviles. En particular, la utilización en estos sistemas de un canal “abierto” entre los equipos de usuario y la infraestructura como es el canal radio, constituye uno de los aspectos fundamentales que condicionan el desarrollo de la arquitectura y mecanismos de seguridad. Además de fuertes condicionantes derivados del uso de una interfaz radio, la gestión de seguridad también debe ocuparse de que las comunicaciones entre los diferentes equipos que conforman la infraestructura del sistema se realice con las garantías de seguridad necesarias. Así por ejemplo, si en un emplazamiento de una estación base LTE se utilizan radioenlaces de microondas para sustentar las interfaces S1 y X2 que conectan la base con el resto de la infraestructura, la información transportada a través de estas interfaces contiene información crítica tal como identidades de usuario, claves de seguridad, etc., que también debe estar debidamente protegida.

En el siguiente apartado se introducen algunos conceptos básicos relativos a la provisión de servicios de seguridad y se describe la arquitectura global de seguridad considerada en el sistema LTE. Sobre esta base, en los apartados posteriores se analizan más detalladamente

los mecanismos de seguridad de acceso a la red (apartado 3.4.2) y los mecanismos de seguridad en la parte de la infraestructura (apartado 3.4.3). Finalmente, en el apartado 3.4.4 se analizan los procedimientos de gestión de seguridad más relevantes.

3.4.1 Marco general de seguridad

Los mecanismos de seguridad que incorporan las redes de comunicaciones proporcionan diferentes servicios de seguridad¹⁶. De forma general, existen cinco tipos de categorías en los que pueden clasificarse los servicios de seguridad en los sistemas de comunicaciones [10]:

- **Autenticación.** Un servicio de autenticación debe garantizar que la entidad con la que establece la relación es quien dice ser. Este servicio es clave en redes móviles para contrastar la identidad de los usuarios que se conectan al sistema, así como para que estos usuarios tengan la certeza que están accediendo a la red correcta (autenticación mutua entre usuarios y red).
- **Control de Acceso.** Este servicio de seguridad se ocupa de la prevención de cualquier acceso no autorizado a los recursos. Así, en el contexto de las redes de comunicaciones móviles, es muy importante condicionar la activación de los servicios de comunicación a los derechos de acceso que tenga el usuario en virtud de su tipo de suscripción al sistema.
- **Confidencialidad.** Este servicio se ocupa de ofrecer la protección de los datos que transporta el sistema frente a observadores no autorizados. El servicio de confidencialidad es uno de los componentes básicos de la interfaz radio en aras a evitar que un receptor cualquiera pueda “entender” la información transmitida por un usuario determinado. La confidencialidad puede aplicarse tanto a la información propia de los servicios que estén utilizando los usuarios (e.g., confidencialidad de los datos contenidos en los paquetes IP) así como a la señalización propia relacionada con la operativa del sistema (e.g., señalización de control RRC en la interfaz radio). Los servicios de confidencialidad también pueden abarcar la protección de la información que puede derivarse de la observación de flujos de tráfico (es decir, aunque no se pueda “entender” lo que transmite un usuario, el hecho de conocer cuándo transmite y dónde también puede ser objeto de protección). Este último caso de confidencialidad, en sistemas LTE suele denominarse como privacidad y, tal como se verá, se basa en limitar al máximo la transmisión de identificadores universales de los usuarios (i.e., IMSI) en la interfaz radio.
- **Integridad.** Mediante este tipo de servicio de seguridad se garantiza que los datos recibidos por una entidad no han sufrido alteración alguna desde su emisión. El servicio de integridad es muy importante para evitar ataques al sistema que consistan en alterar la información que circula por él y causar comportamientos no deseados.
- **No repudio.** Este servicio pretende dejar constancia de la participación de una entidad en un proceso o transferencia de información. Mediante este servicio, el recipiente de

¹⁶ Un servicio de seguridad es aquel servicio de procesado o comunicación destinado a ofrecer algún tipo de protección específica a los recursos del sistema [10][12].

la información dispone de “pruebas” que le permiten identificar el origen de los datos (frente a una posible negación del emisor de los datos como origen de los mismos). El emisor de los datos puede también disponer de “pruebas” para constatar que sus datos han sido recibidos por un destinatario (ante una posible negación del recipiente acerca de la recepción de la información). Los servicios de no repudio se basan principalmente en el uso de mecanismos de firma electrónica.

Tal como se verá en los siguientes apartados, el marco de seguridad de un sistema LTE cubre aspectos de autenticación (más concretamente autenticación mutua), control de acceso (i.e., autorización), confidencialidad e integridad. Por otro lado, el sistema LTE no incorpora mecanismos de no repudio en el marco de seguridad entre usuarios y red ya que su aplicación se fundamenta en la utilización de certificados digitales, opción que no ha sido incorporada en LTE.

El diseño de los diferentes mecanismos de seguridad que incorpora el sistema LTE se ha realizado en base a la definición de una arquitectura completa de seguridad que integra todos los posibles componentes y agentes en la provisión segura de los servicios de telecomunicación. La arquitectura de seguridad de los sistemas 3GPP se muestra en la Figura 3.17.

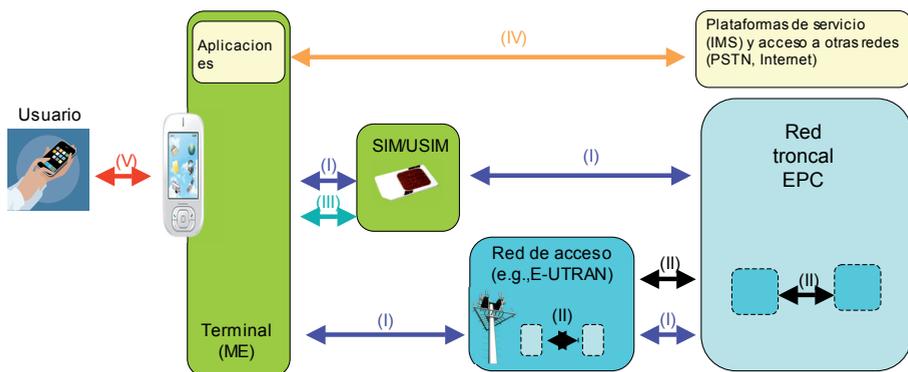


Figura 3.17 Dominios de seguridad (3GPP TS 33.401 [11])

La arquitectura de seguridad, mostrada en la Figura 3.17, se estructura en dominios de seguridad. Cada dominio de seguridad agrupa las diferentes funciones de seguridad que tienen en común un determinado tipo de vulnerabilidades y unos objetivos de seguridad específicos. Los cinco dominios de seguridad identificados en dicha arquitectura son:

- Seguridad de acceso a la red (I). Este grupo de funciones tiene como objetivo proporcionar a los usuarios un acceso seguro a la red LTE. En particular, la seguridad de acceso a la red alberga funciones de autenticación mutua usuario-red, confidencialidad (incluyendo privacidad) e integridad. Los mecanismos de seguridad necesarios para proporcionar estos servicios son específicos de la red de acceso (E-UTRAN, UTRAN o GERAN).

- Seguridad en el dominio de red (II). Funciones de seguridad relacionadas con el soporte de servicios de seguridad entre los equipos que componen la infraestructura de red. Es decir, los intercambios de información entre los equipos de la red de un operador móvil (e.g., interfaces entre eNBs y equipos de la red troncal) así como entre equipos de diferentes operadores (e.g., acceso a la base de datos HSS de un operador matriz para la provisión de un servicio de *roaming*) deben estar adecuadamente protegidos.
- Seguridad del equipo de usuario (III). Este dominio de seguridad abarca aquellas funciones destinadas a establecer un marco de operación seguro entre el terminal LTE, la tarjeta SIM/USIM y el propio usuario del sistema. Un ejemplo claro de estas funciones lo encontramos en la utilización de un número secreto (PIN) para poder operar una tarjeta SIM/USIM en un terminal. El uso de un PIN es la forma más extendida actualmente de vincular al usuario con su módulo de identificación en la red (tarjeta SIM/USIM).
- Seguridad en nivel de aplicación y servicios (IV). Este dominio albergaría las funciones de seguridad utilizadas por las diferentes aplicaciones a las que tiene acceso el usuario a través de la red LTE. Estas funciones aplicarían directamente entre el equipo de usuario y los servidores de las plataformas de servicios o entidades remotas con las que se intercambiara la información. A modo de ejemplo, el servicio de confidencialidad (cifrado) de la señalización SIP que soporta el subsistema IMS formaría parte de este dominio. También, la utilización de protocolos de transporte seguros para el intercambio de información (e.g., *Transport Layer Security, SSL*) formarían parte de este grupo de funciones. Nótese que, de hecho, este nivel de seguridad es transparente a la red LTE ya que, tal como se ha visto, ésta se ocupa de ofrecer un servicio de conectividad IP y no limita el tipo de protección que pueda hacerse a niveles superiores (protocolos de transporte y señalización de aplicación).
- Visibilidad y configuración de la seguridad (V). Aquí se incluyen el conjunto de funciones destinadas a proporcionar al usuario una visión clara de los servicios de seguridad que están operativos en la red. Un ejemplo claro sería la implementación de un indicador gráfico en el equipo de usuario (e.g., un candado cerrado) que informara al usuario que su información se está transfiriendo mediante el uso del servicio de confidencialidad (cifrado) radio que tiene E-UTRAN. Nótese que esta práctica tan habitual en navegadores Web (señalización de conexiones seguras mediante el icono de un candado) no lo es en sistemas de comunicaciones móviles donde, como la activación del modo cifrado es potestad del operador de red, el usuario normalmente no sabe si sus comunicaciones están protegidas o no. Además de la visibilidad de los servicios de seguridad, también es importante poder configurar los servicios de seguridad, es decir, que el usuario pueda decidir en qué condiciones de seguridad quiere ejecutar un servicio.

En los apartados posteriores se analizan más detalladamente los mecanismos de seguridad de acceso a la red (dominio (I)) y los mecanismos de seguridad en la parte de la infraestructura (dominio (II)).

3.4.2 Seguridad de acceso a la red

La seguridad de acceso a la red LTE a través de una red de acceso E-UTRAN se compone de los siguientes elementos:

- Mecanismos para la autenticación mutua entre el usuario y la red. El procedimiento a través del cual se realiza la autenticación mutua, junto con la gestión de claves comentada a continuación, se denomina *EPS Authentication and Key Agreement (AKA)*. Los detalles de este procedimiento se describen más adelante en el apartado 3.4.4.1.
- Mecanismos para la determinación de las claves secretas utilizadas en los algoritmos de cifrado para la provisión de los diferentes servicios de confidencialidad e integridad.
- Servicios de confidencialidad e integridad para la transferencia de la señalización NAS entre el equipo de usuario y la entidad MME de la red troncal EPC.
- Servicios de confidencialidad e integridad para la transferencia de la señalización del protocolo RRC entre el equipo de usuario y el eNB (el cifrado se realiza en la capa PDCP de la torre de protocolos radio).
- Servicios de confidencialidad para la transferencia de información en el plano de usuario entre el equipo de usuario y el eNB (el cifrado se realiza en la capa PDCP de la torre de protocolos radio). Nótese que la información del usuario no dispone de un servicio de integridad ya que, en caso de ser necesario, se considera un aspecto dependiente del servicio final en cuestión.

En la Figura 3.18 se ilustran estos componentes.

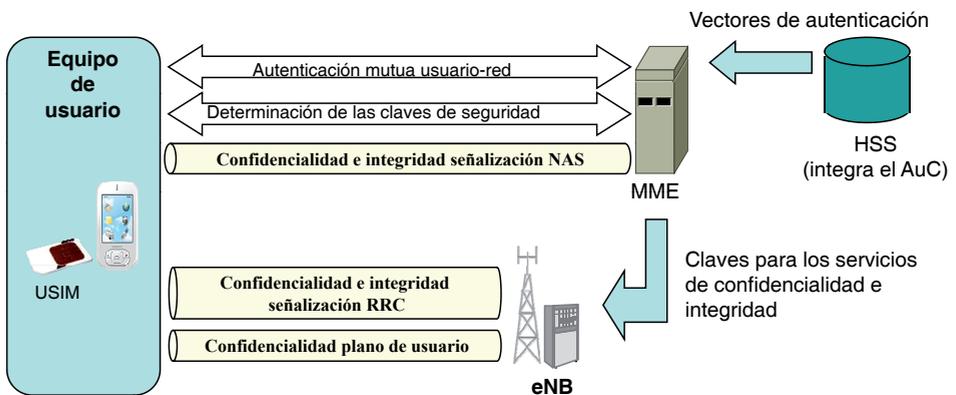


Figura 3.18 Componentes de la seguridad de acceso a la red

Las claves de cifrado y los diferentes parámetros (e.g., *checksums* de seguridad) utilizados en los mecanismos y servicios de seguridad tienen un origen común: una clave secreta **K** asociada a un usuario que se almacena en (1) la tarjeta USIM, y (2) en la entidad funcional

denominada *Authentication Center* (AuC) que forma parte del HSS. A partir de la clave **K** se derivan todas las claves y parámetros utilizados en las funciones de seguridad.

En la infraestructura de red, la clave **K** nunca sale del HSS. En su lugar, el HSS genera los denominados vectores de autenticación. Los vectores de autenticación son la información de seguridad que reciben las entidades MME de cara a sustentar los diferentes procedimientos de seguridad (e.g., EPS AKA), tal como se ilustra en la Figura 3.18. El vector de autenticación utilizado en el sistema LTE se compone de cuatro parámetros:

- **RAND**: Número aleatorio que se utiliza en la generación del resto de parámetros del vector.
- **XRES**: Resultado esperado utilizado en la autenticación del usuario por parte de la red en el mecanismo EPS AKA.
- **AUTN**: Parámetro utilizado para que la USIM del usuario pueda autenticar a la red en el mecanismo EPS AKA.
- K_{ASME}^{17} : Clave de seguridad intermedia utilizada en el sistema LTE para la obtención del conjunto de claves utilizadas en los diferentes servicios de confidencialidad e integridad. Esta clave se deriva de unas claves denominadas CK (*Ciphering Key*) e IK (*Integrity Key*) que son las utilizadas en UMTS. Por este motivo, a partir de un vector de autenticación válido en LTE puede derivarse un vector para UMTS (cinco parámetros en lugar de cuatro) y viceversa.

En la Figura 3.19 se muestra gráficamente el contenido del vector de autenticación en LTE junto con las funciones y parámetros de entrada que se utilizan para su obtención. El resto de parámetros que se muestran en la figura son:

- **IMSI**: Identificador del usuario. Se utiliza en la generación de la clave K_{ASME} .
- **SQN**: Número de secuencia asociado a un IMSI. Se utiliza para disponer de un mecanismo que permita ordenar los vectores de autenticación secuencialmente y evitar, por ejemplo, ataques de repetición (e.g., utilización malintencionada de un vector de autenticación “viejo”).
- **SN-ID**: Identificador de la red que proporciona el servicio de conectividad al usuario. Mediante la inclusión de este parámetro, las claves K_{ASME} generadas en el HSS se particularizan para la red utilizada que puede ser diferente gracias al servicio de itinerancia (*roaming*) y a la posibilidad de acceder a los servicios a través de redes no 3GPP.

Detalles adicionales sobre la generación del vector de autenticación se proporcionan en TS 33.401 [11].

¹⁷ ASME: Access Security Management Entity

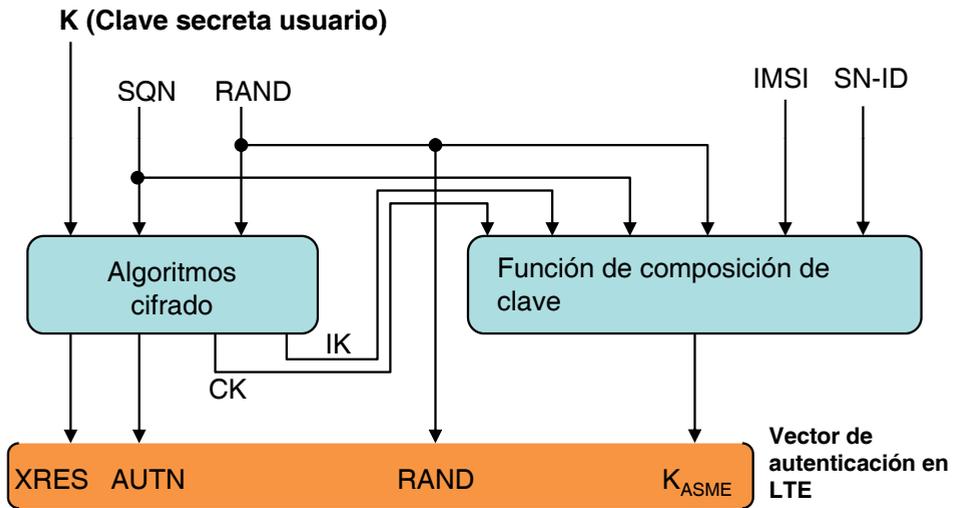


Figura 3.19 Vector de autenticación en LTE

La entidad MME utiliza el vector de autenticación para sustentar el mecanismo EPS AKA (descrito en el apartado 3.4.4.1). Además, a partir de clave K_{ASME} contenida en el vector de autenticación se derivan el resto de claves utilizadas por los diferentes servicios de confidencialidad e integridad. La jerarquía de claves de LTE para la implementación de tales servicios se muestra en la Figura 3.20. Las claves K_{NASenc} y K_{NASint} son utilizadas, respectivamente, en el servicio de confidencialidad e integridad para la transferencia de la señalización NAS. Ambas claves son derivadas a partir de la clave K_{ASME} en el terminal y en la entidad MME, que son los dos puntos de terminación de los protocolos NAS. A partir de la clave K_{ASME} también se deriva una clave de seguridad K_{eNB} que, en la parte de la red, la entidad MME transfiere al eNB al que se encuentra conectado el terminal. De esta forma, tanto eNB como terminal, a partir del conocimiento de la clave K_{eNB} , pueden derivar las claves utilizadas en la interfaz radio para proporcionar confidencialidad al plano de usuario (clave K_{UPenc}) y confidencialidad e integridad a la señalización RRC (claves K_{RRCenc} y K_{RRCint}). La clave K_{eNB} y sus claves derivadas nunca se transmiten por la interfaz radio.

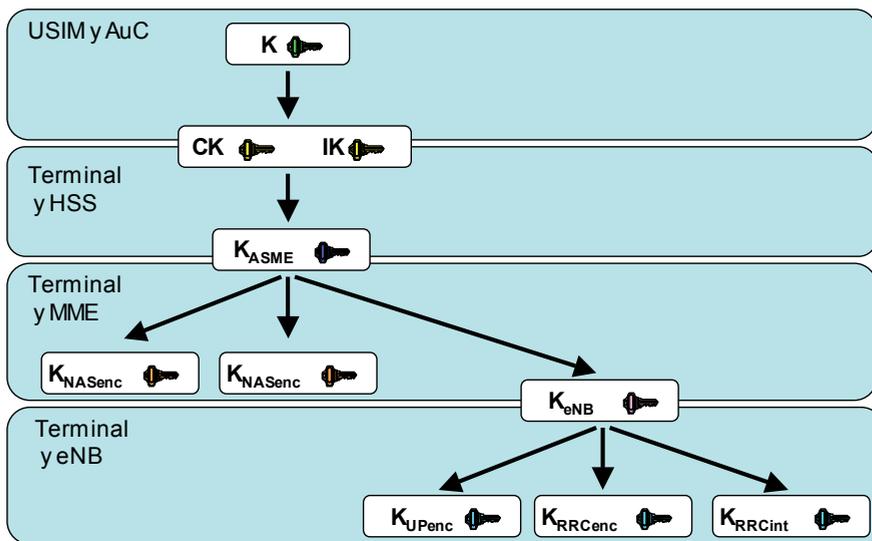


Figura 3.20 Jerarquía de claves de seguridad

Los algoritmos de cifrado para la provisión de los servicios de confidencialidad e integridad en E-UTRAN se denominan UEA2 (UMTS *Encryption Algorithm 2*) y UIA2 (UMTS *Integrity Algorithm 2*), ambos especificados en la Release 7 de UMTS. Como alternativa, E-UTRAN también contempla la utilización de algoritmos basados en AES (*Advanced Encryption Standard*), uno de los algoritmos más populares de cifrado simétrico que existen en la actualidad. Una descripción más detallada de los algoritmos de cifrado e integridad utilizados en los diferentes servicios de seguridad puede encontrarse en TS 33.401 [11].

3.4.3 Seguridad en la infraestructura de red

La solución planteada para proporcionar seguridad en las interfaces internas de la infraestructura de red basadas en torres de protocolos IP se denomina *Network Domain Security for IP (NDS/IP)* y se especifica en el documento 3GPP TS 33.210 [13]. NDS/IP establece los mecanismos de seguridad que pueden utilizarse para la transferencia segura de información entre los equipos que forman parte de la infraestructura de red de un mismo operador así como entre los equipos que enlazan las redes de diferentes operadores. NDS/IP se ha diseñado para proteger los protocolos del plano de control, a través de los cuales se transporta la información más sensible (e.g., transferencia de datos de suscripción, vectores de autenticación, etc.). No obstante, la solución NDS/IP también es válida para proteger los protocolos de plano de usuario en las interfaces internas de la red (e.g., interfaces basadas en GTP-U).

La solución NDS/IP se basa en la utilización de IPsec definido en IETF RFC4301 [14]. IPsec es una colección de protocolos y algoritmos para proporcionar seguridad en la capa de red, incluyendo la gestión de claves. En la solución NDS/IP, el protocolo de seguridad IPsec

utilizado es ESP (*Encapsulating Security Payload*) definido en IETF RFC 4303 [15]. ESP permite ofrecer servicios de confidencialidad, integridad y autenticación del origen de datos. En la Figura 3.21 se ilustra el formato de un paquete del protocolo IPsec ESP operando en modo túnel. En este modo, el protocolo IPsec ESP permite establecer un túnel entre dos puntos de una red IP (análogamente a los protocolos GTP y GRE vistos en el apartado 2.6.3 del Capítulo 2) donde el paquete IP original se transporta encapsulado en otro paquete IP junto con cabeceras especificadas por el protocolo ESP. La dirección IP del paquete original se denomina dirección IP interna del túnel (i.e., *inner address*) y la dirección del paquete ESP en modo túnel se refiere como dirección externa del túnel (i.e., *outer address*). El servicio de confidencialidad cubre el paquete IP original junto con una serie de bits de *padding* (relleno) que pueden ser necesarios en función del algoritmo de cifrado que se utilice (los algoritmos de cifrado suelen procesar la información en bloques de bits de tamaño fijo). Por otro lado, la integridad y autenticación del paquete IP original cifrado y la cabecera ESP se realiza mediante un *checksum* criptográfico que se añade al final del paquete ESP. Además del modo túnel representado en la figura, el protocolo ESP también soporta el modo transporte. A diferencia del modo túnel, el modo transporte no añade una cabecera externa adicional sino que utiliza la cabecera del paquete IP original (con lo que la cabecera no queda protegida por el servicio de confidencialidad ya que ésta debe ser procesada por los *routers* de la red).

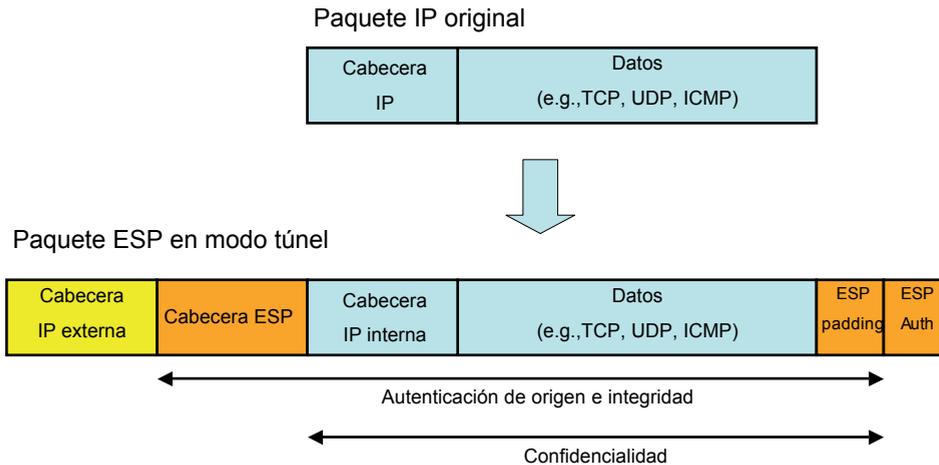


Figura 3.21 Formato de un paquete IPsec ESP en modo túnel

La aplicación de IPsec en la solución NDS/IP es diferente en función de cómo se encuentre estructurada la red en términos de dominios de seguridad. Un dominio de seguridad se concibe el conjunto de equipos que una red que están gestionados por la misma autoridad administrativa. Un ejemplo de un dominio de seguridad podría ser la red de un mismo operador de red (e.g., una red LTE, una red de tránsito, etc.), no quedando excluida la posibilidad de un operador decida estructurar su red en varios dominios de seguridad diferentes. Bajo esta perspectiva, la arquitectura de la solución NDS/IP se representa en la Figura 3.22. Las

entidades de red representadas en la figura pueden ser cualquiera de las entidades de red que forma parte de la arquitectura del sistema LTE (e.g., eNB, MME, P-GW, etc.). El tráfico que se transfiere entre dominios de seguridad diferentes debe protegerse mediante la utilización de pasarelas de seguridad (i.e., SEG). Las pasarelas de seguridad utilizan IPsec ESP en modo túnel para enviar la información entre dominios (e.g., entre las redes del operador matriz y la red visitada en un servicio de itinerancia). Las pasarelas utilizan el protocolo *Internet Key Exchange* (IKE), bien la versión IKEv1 o IKEv2 [16], para establecer las asociaciones de seguridad¹⁸ necesarias para el funcionamiento de IPsec. Por otro lado, en las comunicaciones entre equipos del mismo dominio de seguridad, el operador es quien decide el grado de seguridad que quiera utilizar. En este caso, pueden existir interfaces que no se protejan e interfaces que hagan uso de los mismos mecanismos especificados entre las pasarelas de seguridad. En las transferencias dentro de un dominio de seguridad el protocolo ESP puede utilizarse también en modo transporte. Una interfaz especialmente crítica en la infraestructura de una red LTE es la interfaz S1 entre la red troncal EPC y un eNB. Tal como se ha comentado en el apartado 3.4.2, los servicios de confidencialidad del plano de usuario en LTE únicamente abarcan la interfaz radio (se soportan en la capa PDCP). Por tanto, en el plano de usuario de la interfaz S1, el tráfico IP cursado por los usuarios se transmite sin protección por lo que la aplicación de los mecanismos considerados en la solución NDS/IP puede ser apropiada atendiendo a la seguridad física del emplazamiento donde se encuentre instalado el eNB.

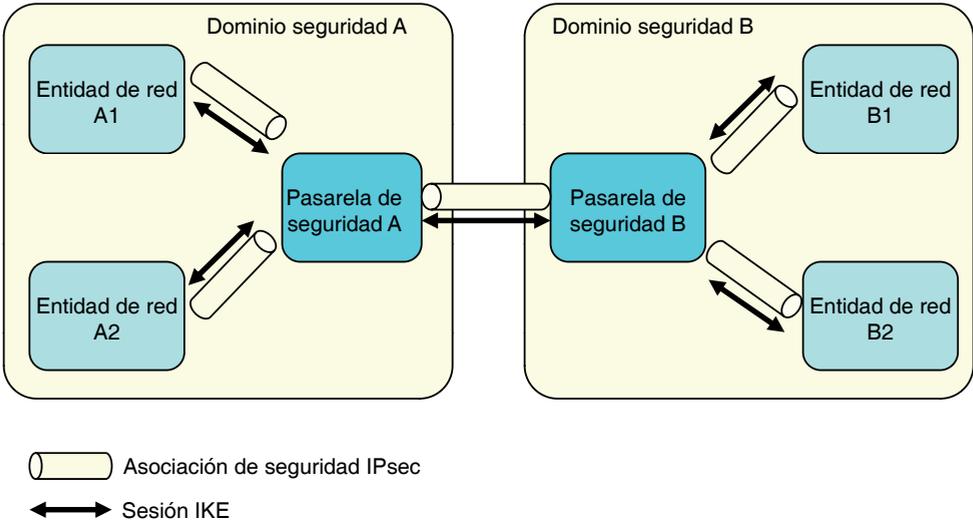


Figura 3.22 Arquitectura de la solución NDS/IP

¹⁸ Una asociación de seguridad (SA, Security Association) consiste en un conjunto de parámetros que comparte las dos entidades de red entre las cuales se establece la relación segura. Una SA indica los algoritmos criptográficos utilizados, las claves de cifrado, la duración de las claves, etc.

3.4.4 Procedimientos de gestión de seguridad

Los principales procedimientos asociados con la gestión de seguridad son los siguientes:

- Autenticación y selección de clave. El mecanismo que se utiliza en el acceso a través de E-UTRAN se denomina EPS AKA y se especifica en TS 33.401 [11].
- Activación del modo seguro para la señalización NAS. Mediante este procedimiento la entidad MME controla la activación y el modo de operación de los servicios de confidencialidad e integridad que se aplican a la señalización NAS.
- Comprobación de la identidad del terminal. Este procedimiento permite interrogar al equipo de usuario desde la entidad MME para que indique la identidad del terminal (i.e., *International Mobile Equipment Identity*, IMEI). Con esta información el operador de la red podría, por ejemplo, detectar la utilización de un terminal robado a través de las comprobación pertinente en la base de datos EIR (*Equipment Identity Register*) considerada en redes 3GPP.
- Activación del modo seguro en la red de acceso E-UTRAN. Mediante este procedimiento la entidad MME indica a un eNB que active los servicios de confidencialidad e integridad. El establecimiento de estos servicios se señala mediante mensajes RRC específicos entre eNB y equipo terminal.

Una descripción detallada del mecanismo de activación del modo seguro para la señalización NAS y de la comprobación de la identidad del terminal puede encontrarse en TS 23.401 [4]. La activación del modo seguro en la interfaz radio se cubre en TS 36.331 [17]. Los detalles principales del procedimiento EPS-AKA se cubren en el siguiente apartado.

3.4.4.1 Authentication and Key Agreement

Mediante el procedimiento EPS AKA el usuario (i.e., la aplicación USIM) y la red LTE se autentican mutuamente. Además, como resultado del procedimiento queda establecida la clave intermedia K_{ASME} a partir de la cual se derivan todas las claves utilizadas en los servicios de confidencialidad e integridad.

El mecanismo EPS AKA asegura que las claves de cifrado e integridad no se han utilizado anteriormente. El mecanismo es compatible con otros sistemas 3G debido a que los principales organismos de estandarización (3GPP, 3GPP2) han adoptado el mismo protocolo de autenticación. El mecanismo EPS AKA puede soportarse con una tarjeta USIM *Release 99* o posterior.

El procedimiento se ilustra en la Figura 3.23 y sus pasos se describen a continuación.

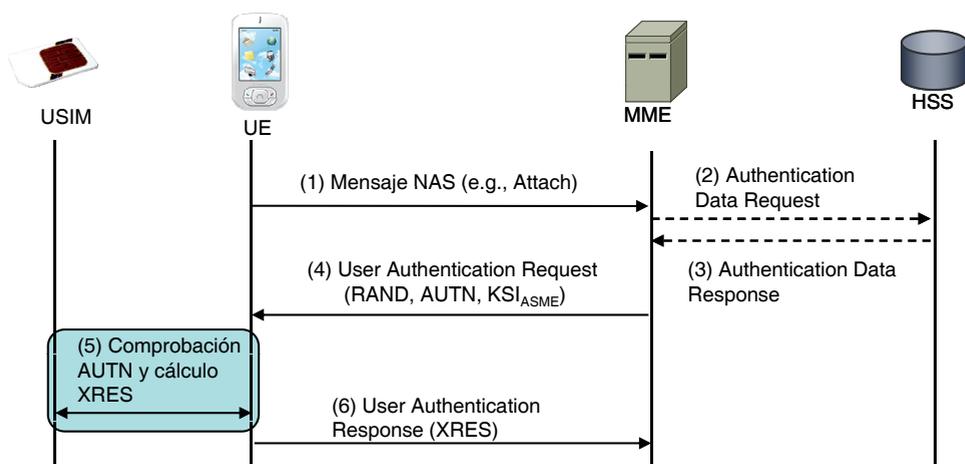


Figura 3.23 Procedimiento EPS AKA

(1) Disparo del procedimiento

El procedimiento EPS AKA puede ejecutarse tras la recepción del primer NAS enviado por el terminal móvil durante el inicio de procedimientos tales como el registro en la red, actualización del área de seguimiento y petición de servicio.

(2) y (3) Obtención de vectores de autenticación

En caso de que la entidad MME no disponga de vectores de autenticación del usuario, ésta los solicita a la base de datos HSS. En la solicitud de los vectores la entidad MME debe indicar el IMSI del usuario y el tipo de tecnología de acceso que se utiliza (los vectores de autenticación servidor por la base de datos HSS son diferentes para, por ejemplo, E-UTRAN o UTRAN). El contenido del vector de autenticación para accesos a través de E-UTRAN se detalla en el apartado 3.4.2.

(4) Petición de autenticación al usuario

El mensaje enviado por la entidad MME contiene el número aleatorio RAND y el parámetro AUTN (*Authentication Token*), dos de los cuatro parámetros que forman el vector de autenticación. El mensaje también contiene un identificador corto KSI_{ASME} que se asocia a la clave K_{ASME} para su identificación en otros mecanismos donde se tenga que proporcionar una referencia a ella.

(5) Comprobación AUTN y generación XRES

La tarjeta USIM comprueba la validez del parámetro AUTN que le permite autenticar la identidad de la red con la que está manteniendo el procedimiento EPS AKA. Las comprobaciones realizadas se especifican en TS 33.102 [18]. Si todo es correcto, la USIM genera el parámetro de respuesta XRES junto con las dos claves CK y IK a partir de las cuales el terminal obtendrá la clave intermedia K_{ASME}.

(6) Respuesta de autenticación

El terminal responde con el parámetro XRES proporcionado por la USIM. En la entidad MME se comprueba que el parámetro recibido concuerde con el contenido en el vector de autenticación. En caso de que coincidan, el usuario se considera autenticado.

3.5 Referencias bibliográficas

- [1] 3GPP TS 23.203, "Policy and charging control architecture".
- [2] "Quality of Service in IP Networks", Grenville Armitage, Publisher: New Riders, First Edition, April 07, 2000, ISBN: 1-57870-189-9.
- [3] 3GPP TS 23.107, "Quality of Service (QoS) concept and architecture".
- [4] 3GPP TS 23.401, "General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access".
- [5] 3GPP TS 23.003, "Numbering, addressing and identification".
- [6] 3GPP 36.304, "User Equipment (UE) Procedures in *Idle Mode*".
- [7] 3GPP TS 23.122, "Non-Access-Stratum (NAS) functions related to Mobile Station (MS) in *idle mode*".
- [8] 3GPP TR 25.913, "Requirements for evolved UTRA and evolved UTRAN".
- [9] 3GPP TS 36.133, "Evolved Universal Terrestrial Radio Access (E-UTRA); Requirements for support of radio resource management".
- [10] ITU Recommendation X.800, "Security architecture for open systems interconnection for CCITT applications".
- [11] 3GPP TS 33.401, "3GPP System Architecture Evolution (SAE), Security architecture".
- [12] William Stallings, "Cryptography and Network Security", Third Edition, Ed. Prentice Hall, 2003.
- [13] 3GPP TS 33.210, "3G Security; Network Domain Security; IP network layer security".
- [14] IETF RFC 4301, "Security Architecture for the Internet Protocol".
- [15] IETF RFC 4303, "IP Encapsulating Security Payload (ESP)".
- [16] IETF RFC 4306, "Internet Key Exchange Protocol Version 2".
- [17] 3GPP TS 33.331, "Evolved Universal Terrestrial Radio Access (E-UTRA), Radio Resource Control (RRC); Protocol specification (Release 8)".
- [18] 3GPP TS 33.102, "3G security; Security architecture".



IV

Tecnologías de Nivel Físico

4. Tecnologías de nivel físico

Este capítulo presenta los fundamentos de las tecnologías de nivel físico empleadas en el futuro sistema LTE, y que constituyen una de las principales diferencias en relación a los sistemas predecesores de comunicaciones móviles, en tanto que permiten conseguir mayores niveles de capacidad y eficiencia en el uso de los recursos radio. En primer lugar, dentro del capítulo se aborda una descripción de las técnicas de acceso múltiple, OFDMA (*Orthogonal Frequency Division Multiple Access*) y SC-FDMA (*Single Carrier Frequency Division Multiple Access*), empleadas para los enlaces descendente y ascendente, respectivamente. En particular, tras una primera descripción en el apartado 4.1 de los fundamentos de la técnica de transmisión OFDM, en que las nuevas técnicas se sustentan, el apartado 4.2 aborda las características de OFDMA y el apartado 4.3 describe los aspectos fundamentales de SC-FDMA. Tras la descripción de las técnicas de acceso, en el apartado 4.4 se abordan las estructuras de transmisión y recepción con múltiples antenas, presentando los aspectos fundamentales de estas estructuras que permiten lograr significativos aumentos de capacidad. Finalmente, el apartado 4.5 presenta la estructura combinada de transmisores/receptores OFDMA con múltiples antenas.

4.1 Fundamentos de OFDM

La técnica de transmisión OFDM (Orthogonal Frequency Division Multiplex) constituye un mecanismo de transmisión multi-portadora consistente en multiplexar un conjunto de símbolos sobre un conjunto de subportadoras. Gracias a las propiedades de ortogonalidad de dichas subportadoras, es posible efectuar la transmisión simultánea de todos los símbolos manteniendo la capacidad de separación de los mismos en recepción.

Si bien esta técnica es ampliamente conocida desde los años 60, su aplicación práctica en el ámbito de las comunicaciones inalámbricas es mucho más reciente, principalmente debido a la complejidad que involucraba en los equipos transmisores y receptores. Hoy en día es utilizada por sistemas tales como la Televisión Digital Terrestre según el estándar DVB-T o las redes inalámbricas de área local según los estándares IEEE 802.11a/g, a la vez que constituye la base para la técnica de acceso múltiple OFDMA (*Orthogonal Frequency Division Multiple Access*) empleada por el sistema LTE.

La característica fundamental de la técnica OFDM es el empleo de un conjunto de K subportadoras que presentan la propiedad de ser ortogonales. Asumiendo la notación de señales complejas (en el Anexo 4.1 se presenta un pequeño resumen de dicha notación), dichas subportadoras pueden formularse en banda base como:

$$x_k(t) = e^{j2\pi k\Delta f t} \text{rect}_{T_s}(t) \quad 0 \leq k \leq K-1 \quad (4.1)$$

donde $f_k = k\Delta f$ es la frecuencia de la subportadora k -ésima y $\text{rect}_{T_s}(t)$ representa un pulso rectangular con duración entre 0 y T_s . Por otra parte, $\Delta f = 1/T_s$ es la separación entre

subportadoras. Obsérvese que, con esta definición, dos subportadoras diferentes $x_m(t)$ y $x_k(t)$ cumplen la condición de ser ortogonales en el intervalo temporal T_s , lo que significa que la integración del producto de las mismas en dicho intervalo es nula excepto cuando $m=k$:

$$R_{x_m, x_k}(t) = \frac{1}{T_s} \int_0^{T_s} x_m(t) x_k^*(t) dt = \frac{1}{T_s} \int_0^{T_s} e^{j2\pi(m-k)\Delta f t} dt = \frac{1}{T_s} \int_0^{T_s} e^{j2\pi(m-k)\frac{t}{T_s}} dt = \begin{cases} 1 & \text{si } m = k \\ 0 & \text{si } m \neq k \end{cases} \quad (4.2)$$

Nótese igualmente que la propiedad de ortogonalidad entre subportadoras se consigue precisamente gracias a la relación existente entre la separación de las mismas Δf y su duración temporal T_s .

La Figura 4.1 muestra un ejemplo con el módulo de los espectros correspondientes a un conjunto de 6 subportadoras OFDM. Obsérvese en la figura que para cada frecuencia múltiplo de $1/T_s$ únicamente existe contribución espectral de una de las subportadoras, mientras que el resto presentan nulos. A su vez, la Figura 4.2 muestra la evolución temporal de la parte real de las 6 subportadoras.

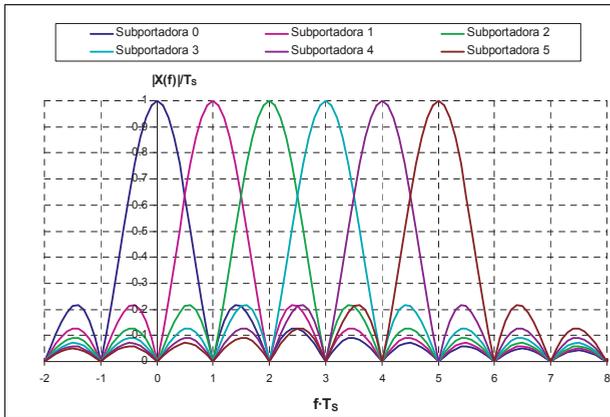


Figura 4.1 Ejemplo del espectro correspondiente a 6 subportadoras OFDM

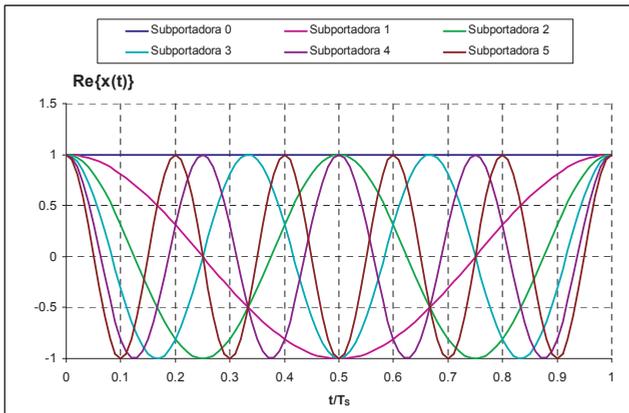


Figura 4.2 Ejemplo de la señal temporal correspondiente a 6 subportadoras OFDM

4.1.1 Esquemas de modulación y demodulación

Con objeto de estudiar el principio de modulación de una señal mediante OFDM, considérese que se dispone de un conjunto de K símbolos complejos d_0, d_1, \dots, d_K que se desean transmitir simultáneamente, siendo T_s la duración de un símbolo, también denominada período de símbolo. Cada uno de dichos símbolos se modula mediante una de las subportadoras definidas en (4.1), tal y como se muestra en la Figura 4.3. La señal resultante del proceso, denominada símbolo OFDM, vendrá dada por:

$$s(t) = \sum_{k=0}^{K-1} d_k e^{j2\pi k \Delta f t} \text{rect}_{T_s}(t) \quad (4.3)$$

Obsérvese que esta señal corresponde a la señal modulada en banda base. Habitualmente en un sistema de comunicaciones radio la señal $s(t)$ se trasladará a continuación a una frecuencia portadora f_0 a la que finalmente se efectuará la transmisión, dando lugar a la señal:

$$r(t) = \sum_{k=0}^{K-1} d_k e^{j2\pi(f_0+k\Delta f)t} \text{rect}_{T_s}(t) \quad (4.4)$$

en consecuencia, la frecuencia de las subportadoras finalmente transmitidas pasa a ser $f_0+k\Delta f, 0 \leq k \leq K-1$.

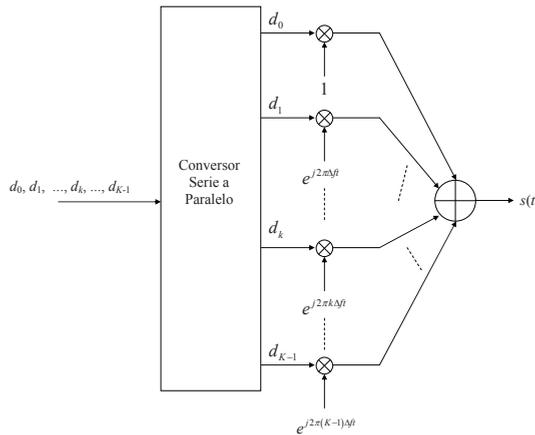


Figura 4.3 Modulación OFDM de un conjunto de símbolos

Como resultado del proceso de modulación explicado, el espectro de la señal banda base $s(t)$ viene dado por:

$$S(f) = e^{-j\pi T_s f} T_s \sum_{k=0}^{K-1} d_k \frac{\sin(\pi(f-k\Delta f)T_s)}{\pi(f-k\Delta f)T_s} \quad (4.5)$$

A modo de ejemplo, la Figura 4.4 muestra la representación de dicho espectro (sin considerar el término de fase lineal $e^{-j\pi T_s f}$) en el caso de una modulación OFDM con 6 subporta-

doras y la secuencia de símbolos {2,1,-2,3,1,-1}. A su vez, la correspondiente evolución temporal de la amplitud de la señal modulada (símbolo OFDM) se representa en la Figura 4.5.

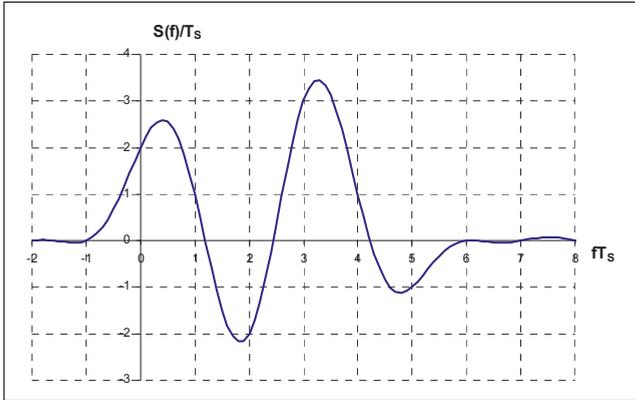


Figura 4.4 Espectro de la señal modulada

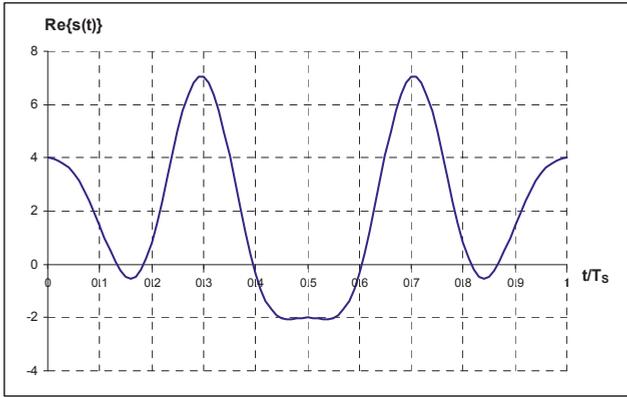


Figura 4.5 Evolución temporal del símbolo OFDM

En recepción, tras efectuar la traslación espectral de la señal $r(f)$ a banda base, dando lugar de nuevo a la señal $s(t)$, el proceso de demodulación se muestra en la Figura 4.6. Consta de K ramas diferentes, cada una formada por una multiplicación compleja por la correspondiente subportadora seguida de una integración en el período de símbolo. Analizando la salida de la rama k -ésima se obtiene que:

$$\tilde{d}_k = \sum_{m=0}^{K-1} d_m \frac{1}{T_S} \int_0^{T_S} e^{j2\pi(m-k)\Delta f t} dt = d_k \tag{4.6}$$

Nótese en esta última expresión que, gracias a la ortogonalidad entre subportadoras, la única contribución no nula del sumatorio es la correspondiente a $m=k$, lo que permite obtener a la salida de esta rama el símbolo d_k que se envió sobre la subportadora k -ésima.

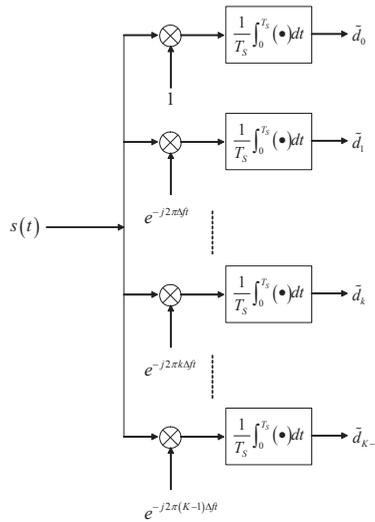


Figura 4.6 Demodulación OFDM

Si bien los esquemas presentados en las anteriores figuras han resultado de utilidad para explicar el principio de funcionamiento de la modulación y la demodulación de señales OFDM, es habitual realizar la implementación práctica de los moduladores y demoduladores mediante técnicas basadas en la transformada discreta de Fourier (DFT- *Discrete Fourier Transform*), tal y como se explica a continuación.

Dada una señal $y(t)$ muestreada en el dominio temporal con N muestras $y(n)$, $n=0, \dots, N-1$, su DFT viene definida como la secuencia de N muestras:

$$Y(k) = \sum_{n=0}^{N-1} y(n) e^{-j\frac{2\pi}{N}kn} \quad k = 0, 1, \dots, N-1 \quad (4.7)$$

A su vez, la correspondiente transformada discreta de Fourier inversa (IDFT - Inverse Discrete Fourier Transform) proporciona las muestras de la señal $y(n)$ como:

$$y(n) = \frac{1}{N} \sum_{k=0}^{N-1} Y(k) e^{j\frac{2\pi}{N}kn} \quad n = 0, 1, \dots, N-1 \quad (4.8)$$

Por otro lado, si analizamos la expresión de la señal modulada OFDM y la muestreamos con frecuencia de muestreo $f_m = N\Delta f$ y período $T_m = 1/(N\Delta f) = T_s/N$ (esto es, tomando N muestras de señal en cada período de símbolo T_s), obtenemos:

$$s(n) = \sum_{k=0}^{K-1} d_k e^{j2\pi k\Delta f n T_m} \text{rect}_{T_s}(nT_m) = \sum_{k=0}^{K-1} d_k e^{j\frac{2\pi kn}{N}} = \sum_{k=0}^{N-1} S(k) e^{j\frac{2\pi kn}{N}} \quad n = 0, 1, \dots, N-1 \quad (4.9)$$

siendo la secuencia $S(k)$ una versión extendida de d_k construida como $S(k) = d_k$ para $0 \leq k < K$ y $S(k) = 0$ para $K \leq k < N$. Por lo tanto, como puede verse comparando esta última expresión (4.8)

y prescindiendo del factor multiplicativo constante ($1/N$), las muestras de la señal modulada $s(n)$ no son más que la IDFT de la secuencia $S(k)$ que contiene en las K primeras muestras los símbolos a modular y ceros en las $N-K$ muestras restantes.

En el dominio de la frecuencia, la secuencia $S(k)$ se corresponde con la DFT de la secuencia $s(n)$, y está compuesta por N muestras equiespaciadas en el intervalo 0 a f_m y por lo tanto con una separación igual a la separación entre subportadoras $\Delta f = f_m/N$. Para ilustrar esto, obsérvese el ejemplo del espectro de la señal OFDM mostrada en la Figura 4.4, correspondiente a la modulación de la secuencia de 6 símbolos $\{2,1,-2,3,1,-1\}$. Puede apreciarse como los valores de los 6 símbolos se corresponden con las muestras del espectro tomadas cada $\Delta f = 1/T_s$, tal y como se ilustra en la Figura 4.7, es decir, los símbolos a transmitir no son más que los valores de la DFT de la señal modulada.

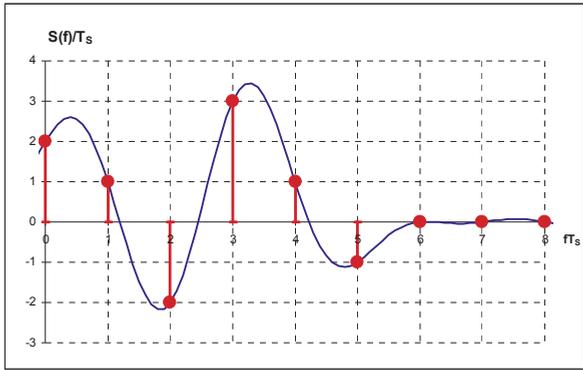


Figura 4.7 Espectro muestreado de una señal OFDM

De acuerdo con todo lo anterior, la forma más habitual de generar una modulación OFDM es considerar los K símbolos a modular como los valores de las muestras en frecuencia de la señal resultante (a razón de una muestra por subportadora) y a partir de dichos símbolos calcular la IDFT de N muestras. Este proceso proporcionará las muestras temporales de la señal que finalmente se pasarán por un convertor Digital/Analógico (D/A) para generar la señal analógica a transmitir, tal y como se muestra en la Figura 4.8.

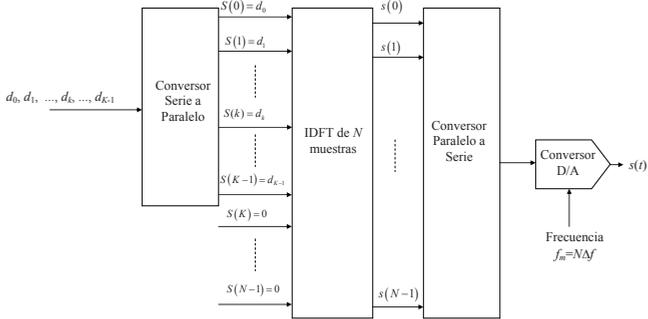


Figura 4.8 Modulación OFDM basada en IDFT.

Hay que destacar que la frecuencia de muestreo viene dada por el número de muestras de la DFT, N , y por la separación entre subportadoras, Δf , como $f_m = N\Delta f$. Por otra parte, el ancho de banda de la señal transmitida viene determinado por el número de subportadoras ocupadas, K , siendo aproximadamente $K\Delta f$.

Por otro lado, en la implementación del modulador basada en IDFT es habitual emplear la denominada transformada rápida de Fourier (IFFT - *Inverse Fast Fourier Transform* y FFT - *Fast Fourier Transform*) con objeto de acelerar el proceso de cálculo. En estas circunstancias, el valor de N debe siempre tomarse como una potencia de 2.

En recepción, el proceso de demodulación se efectúa a partir de la DFT (o su versión rápida FFT), en tanto que lo que se desea obtener a la salida no es más que las muestras frecuenciales de la señal modulada, que coinciden, tal y como se ha dicho, con los diferentes símbolos enviados. El esquema es el que se muestra en la Figura 4.9. Como puede apreciarse, tras efectuar la DFT, de las N muestras resultantes a la salida, las K primeras constituyen los símbolos que se desean recuperar mientras que las restantes hasta completar las N muestras son descartadas, en tanto que corresponderían con los $N-K$ ceros que se han añadido en emisión para realizar la IDFT.

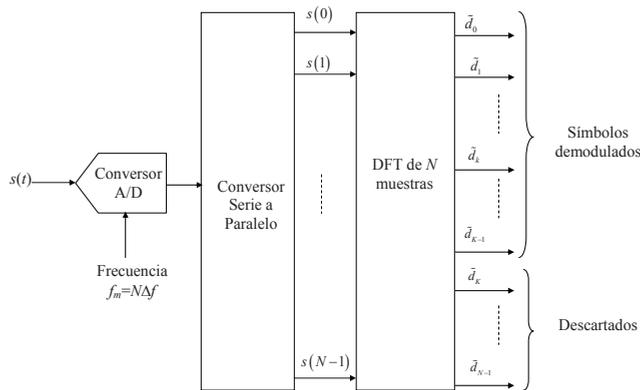


Figura 4.9 Demodulación OFDM basada en DFT

4.1.2 Efecto de la dispersión del canal: concepto de prefijo cíclico

En cualquier sistema de comunicaciones móviles, la propagación se ve afectada por diferentes fenómenos que varían según el entorno del terminal móvil receptor, tales como la presencia de múltiples reflexiones, la difracción por obstáculos, etc. Como resultado de dichos fenómenos, la señal recibida realmente está compuesta por diferentes réplicas de la señal originalmente transmitida, cada una correspondiente a un camino de propagación diferente, y que por lo tanto serán recibidas en diferentes instantes de tiempo y con diferentes amplitudes y fases. El modelo de esta propagación multicamino acostumbra a ser una respuesta

impulsional compuesta por un cierto número de caminos, cada uno caracterizado por un retardo de propagación τ_n y por una amplitud h_n que dependen de los recorridos efectuados por las ondas radioeléctricas en su propagación. Tanto las amplitudes h_n como los retardos τ_n y el número de caminos variarán además temporalmente de acuerdo con las variaciones del entorno y la movilidad asociadas.

Con objeto de interpretar el efecto de la propagación multicamino sobre una señal OFDM, considérese el envío continuado de un conjunto de símbolos OFDM, cada uno con una duración T_S y compuesto a su vez por K subportadoras, lo que corresponde a la siguiente señal en banda base:

$$s_T(t) = \sum_{i=-\infty}^{\infty} s^{(i)}(t - iT_S) = \sum_{i=-\infty}^{\infty} \sum_{k=0}^{K-1} d_k^{(i)} e^{j2\pi k \Delta f t} \text{rect}_{T_S}(t - iT_S) \quad (4.10)$$

En esta última expresión se denota al símbolo OFDM enviado en el intervalo temporal $[iT_S, (i+1)T_S]$ como $s^{(i)}$, que a su vez está compuesto por K símbolos $d_k^{(i)}$, uno por subportadora, de acuerdo con la expresión .

Por otro lado, modelaremos la propagación multicamino típica de los entornos de comunicaciones móviles de acuerdo con la siguiente respuesta impulsional, en la que prescindiremos de la dependencia temporal con t bajo la consideración de que dichas variaciones temporales se dan típicamente en una escala de tiempo mucho mayor que el período de símbolo T_S y por lo tanto se puede suponer que el canal no cambia a lo largo de un símbolo, esto es:

$$h(\tau, t) = \sum_{n=0}^{M(t)-1} h_n(t) \delta(\tau - \tau_n(t)) \approx \sum_{n=0}^{M-1} h_n \delta(\tau - \tau_n) = h(\tau) \quad (4.11)$$

donde M representa el número de caminos detectados, mientras que τ_n y h_n son respectivamente el retardo y la amplitud asociada al camino n -ésimo.

Por otra parte, por simplicidad y sin pérdida de generalidad supondremos $\tau_0=0$ y $h_0=1$, de modo que los diferentes caminos de propagación se definen en relación al primero de ellos $n=0$.

Así, como resultado de la propagación multicamino, la señal recibida, cuyas réplicas temporales se ilustran en la Figura 4.10, se puede expresar matemáticamente como:

$$s_R(t) = \sum_{n=0}^{M-1} h_n s_T(t - \tau_n) = \sum_{n=0}^{M-1} h_n \sum_{i=-\infty}^{\infty} \sum_{k=0}^{K-1} d_k^{(i)} e^{j2\pi k \Delta f (t - \tau_n)} \text{rect}_{T_S}(t - \tau_n - iT_S) \quad (4.12)$$

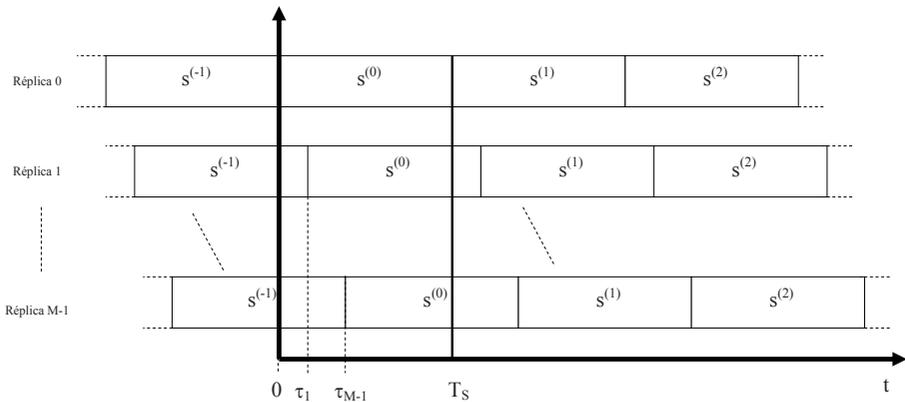


Figura 4.10 Señal recibida tras pasar por un canal con M caminos de propagación

Considérese que la señal recibida se demodula de acuerdo con el esquema presentado anteriormente en la Figura 4.6, analizando la salida de la rama m -ésima en el instante T_S , lo que correspondería a recuperar el símbolo $d_m^{(0)}$ asociado a la subportadora m durante el intervalo $[0, T_S]$. La señal resultante de este proceso de demodulación vendría dada por:

$$\begin{aligned}
 \tilde{d}_m^{(0)} &= \frac{1}{T_S} \int_0^{T_S} s_R(t) e^{-j2\pi m \Delta f t} dt = \\
 &= d_m^{(0)} + \frac{1}{T_S} \sum_{k=0}^{K-1} \sum_{n=1}^{M-1} h_n e^{-j2\pi k \Delta f \tau_n} \left(d_k^{(-1)} \int_0^{\tau_n} e^{j2\pi(k-m)\Delta f t} dt + d_k^{(0)} \int_{\tau_n}^{T_S} e^{j2\pi(k-m)\Delta f t} dt \right) = \\
 &= d_m^{(0)} + \frac{1}{T_S} \sum_{n=1}^{M-1} h_n e^{-j2\pi m \Delta f \tau_n} \left(d_m^{(-1)} \tau_n + d_m^{(0)} (T_S - \tau_n) \right) \\
 &\quad + \frac{1}{T_S} \sum_{\substack{k=0 \\ k \neq m}}^{K-1} \sum_{n=1}^{M-1} h_n e^{-j2\pi k \Delta f \tau_n} \left(d_k^{(-1)} \int_0^{\tau_n} e^{j2\pi(k-m)\Delta f t} dt + d_k^{(0)} \int_{\tau_n}^{T_S} e^{j2\pi(k-m)\Delta f t} dt \right)
 \end{aligned} \tag{4.13}$$

En esta última expresión, notar como, al efectuar el proceso de integración en el intervalo $[0, T_S]$, tal y como se aprecia en la Figura 4.10, únicamente para la primera réplica $n=0$ el período de integración coincide exactamente con el símbolo OFDM $s^{(0)}$ que se debería obtener, mientras que para el resto de réplicas se está integrando parte del símbolo $s^{(0)}$ y parte del símbolo anterior $s^{(-1)}$. Como resultado final de este proceso, tal y como se muestra en la expresión 4.13, a la salida de la demodulación obtenemos, además del término deseado $d_m^{(0)}$, proveniente de la integración de la réplica $n=0$, los siguientes componentes de interferencia:

- interferencia intersimbólica asociada con las diferentes réplicas de la portadora m -ésima, que se aprecia en el segundo sumando de $\tilde{d}_m^{(0)}$, y que hace que la salida dependa del símbolo anterior $d_m^{(-1)}$
- interferencia entre subportadoras, que se aprecia en el tercer sumando de $\tilde{d}_m^{(0)}$, y que

hace que la salida dependa de los símbolos enviados en las subportadoras diferentes de la demodulada ($k \neq m$). En definitiva, esto pone de manifiesto que uno de los efectos de la propagación multicamino sobre una señal OFDM es la pérdida de ortogonalidad entre las diferentes subportadoras.

Estos dos efectos negativos son tanto más significativos cuanto mayor sea el retardo asociado a las diferentes réplicas del canal τ_n en relación al período de símbolo T_s , esto es, cuanto más dispersivo sea el canal.

Con objeto de combatir estos dos efectos negativos de la propagación multicamino, la solución utilizada en OFDM es el empleo en transmisión del denominado prefijo cíclico. Consiste en alargar la transmisión de cada símbolo OFDM hasta una duración total de $T_p + T_s$ a base de repetir, al principio del símbolo, la señal que se envía durante los últimos T_p segundos del símbolo, tal y como se ilustra en la Figura 4.11 a partir del símbolo OFDM que anteriormente se presentó en el ejemplo de la Figura 4.5. En el ejemplo aquí mostrado se ha supuesto una duración del prefijo cíclico igual a la décima parte del período de símbolo.

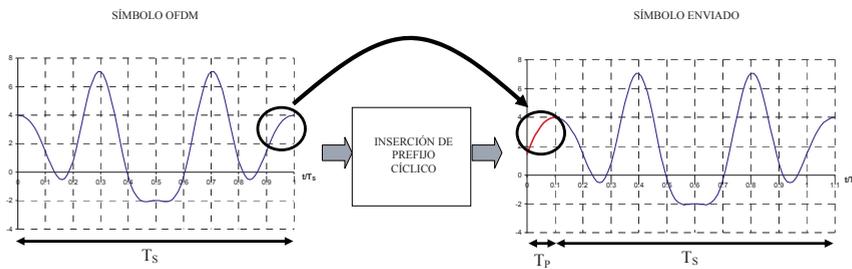


Figura 4.11 Inserción del prefijo cíclico

En recepción, durante el proceso de demodulación la integración se efectuará solamente en el intervalo de duración T_s sin considerar el prefijo cíclico, tal y como se muestra en la Figura 4.12 para una situación con M caminos de propagación. De este modo, en tanto que se escoja una duración de prefijo cíclico superior al máximo retardo de las componentes multicamino ($T_p > \tau_{M-1}$ en el ejemplo), se consigue eliminar totalmente la interferencia intersimbólica. De la misma manera, al repetir en el prefijo cíclico la parte final del símbolo, conseguimos que la integración de cada una de las réplicas se haga sobre el símbolo OFDM completo (observar en la figura que justamente la parte de símbolo que cae fuera del intervalo de integración coincide con la añadida en el prefijo cíclico). De este modo, aseguramos que los productos de subportadoras diferentes se integren en intervalos completos de T_s , lo que permite recuperar la ortogonalidad entre subportadoras y evitar la interferencia entre ellas.

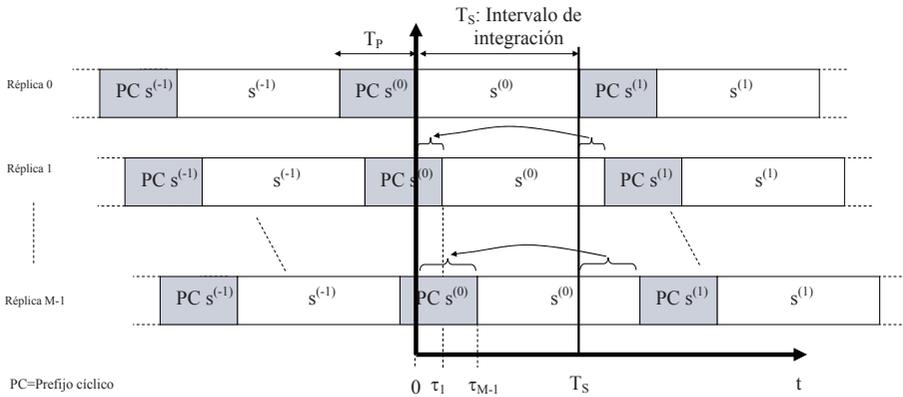


Figura 4.12 Señal recibida e intervalo de integración al utilizar prefijo cíclico

Desde un punto de vista práctico, y teniendo en cuenta la implementación basada en IFFT del modulador OFDM, la inserción del prefijo cíclico se efectúa simplemente repitiendo al principio las muestras finales resultantes del proceso de modulación, lo cual resulta en una implementación muy sencilla.

Cabe destacar en todo caso que el inconveniente de la inserción del prefijo cíclico radica en una menor eficiencia en términos de la potencia transmitida y de la velocidad de transmisión, ya que únicamente una fracción $T_s/(T_p+T_s)$ de la potencia se destina a la parte útil de los símbolos. Análogamente, desde la perspectiva de velocidad de transmisión, la inserción del prefijo cíclico se traduce en una separación mayor entre símbolos consecutivos, de modo que para un espaciado entre subportadoras de $1/T_s$ la tasa de transmisión es de únicamente $1/(T_p+T_s)$ símbolos por portadora en lugar del valor de $1/T_s$ que habría de no usarse prefijo cíclico. De este modo, teniendo en cuenta estas ineficiencias, es preciso escoger el valor de la duración del prefijo cíclico como el mínimo valor posible que permite hacer frente a la dispersión del canal.

Desde un punto de vista matemático, la señal transmitida correspondiente al símbolo OFDM $s^{(0)}$ incluyendo el prefijo cíclico (considerando que $t=0$ se sitúa en el inicio de la parte útil del símbolo mientras que el prefijo cíclico va de $t=-T_p$ a $t=0$) puede expresarse como:

$$\begin{aligned}
 s^{(0)}(t) &= \sum_{k=0}^{K-1} d_k^{(0)} e^{j2\pi k \Delta f t} \text{rect}_{T_s}(t) + \sum_{k=0}^{K-1} d_k^{(0)} e^{j2\pi k \Delta f (t+T_s)} \text{rect}_{T_p}(t+T_p) = \\
 &= \sum_{k=0}^{K-1} d_k^{(0)} e^{j2\pi k \Delta f t} \text{rect}_{T_s}(t) + \sum_{k=0}^{K-1} d_k^{(0)} e^{j2\pi k \Delta f t} \text{rect}_{T_p}(t+T_p)
 \end{aligned}
 \tag{4.14}$$

De este modo, al efectuar la demodulación correspondiente a la subportadora m del símbolo OFDM $s^{(0)}$, realizando la integración entre 0 y T_s sobre la señal multicamino recibida, tal y como se muestra en la Figura 4.12, se obtiene:

$$\begin{aligned} \tilde{d}_m^{(0)} &= d_m^{(0)} + \frac{1}{T_S} \sum_{k=0}^{K-1} \sum_{n=1}^{M-1} h_n e^{-j2\pi k \Delta f \tau_n} \left(d_k^{(0)} \int_0^{\tau_n} e^{j2\pi(k-m)\Delta f t} dt + d_k^{(0)} \int_{\tau_n}^{T_S} e^{j2\pi(k-m)\Delta f t} dt \right) = \\ &= d_m^{(0)} + \frac{1}{T_S} \sum_{k=0}^{K-1} \sum_{n=1}^{M-1} h_n e^{-j2\pi k \Delta f \tau_n} d_k^{(0)} \int_0^{T_S} e^{j2\pi(k-m)\Delta f t} dt = d_m^{(0)} \sum_{n=0}^{M-1} h_n e^{-j2\pi m \Delta f \tau_n} \end{aligned} \quad (4.15)$$

En esta última expresión puede comprobarse como, efectivamente, tanto la interferencia intersimbólica como la interferencia entre subportadoras se han eliminado. Por otro lado, es importante remarcar que a la salida de la rama m -ésima se obtiene el símbolo $d_m^{(0)}$ multiplicado por un factor que depende de las amplitudes y retardos de las diferentes componentes multicamino. Este factor puede relacionarse de forma sencilla con la respuesta frecuencial del canal $H(f)$ que se obtiene como la transformada de Fourier de la respuesta impulsional $h(\tau)$ de , esto es:

$$H(f) = \sum_{n=0}^{M-1} h_n e^{-j2\pi f \tau_n} \quad (4.16)$$

Efectivamente, de acuerdo con , el término multiplicativo a la salida no es más que el valor de la respuesta frecuencial del canal para la frecuencia $m\Delta f$ correspondiente a la subportadora m -ésima, esto es:

$$\tilde{d}_m^{(0)} = d_m^{(0)} H(m\Delta f) \quad (4.17)$$

Por lo tanto, el proceso de modulación y demodulación OFDM incluyendo prefijo cíclico puede modelarse como un conjunto de K canales en paralelo, cada uno asociado a una subportadora y en el que se transmite un símbolo, y que se ve multiplicado por la respuesta frecuencial del canal en dicha subportadora, tal y como se ilustra en la Figura 4.13.

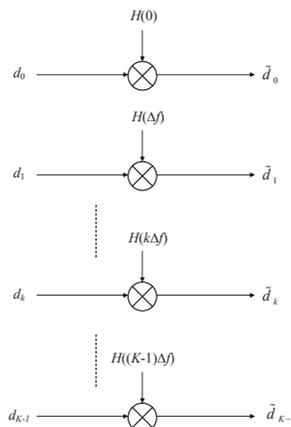


Figura 4.13 Modelo de la transmisión y recepción OFDM como K canales paralelos

En definitiva, un canal con dispersión y por lo tanto selectivo en frecuencia (esto es, una transmisión de banda ancha), gracias a la utilización del prefijo cíclico, se ha transformado en un conjunto de K canales en paralelo, no selectivos en frecuencia (esto es, K transmisiones de banda estrecha). Esto además sugiere la posibilidad de modificar independientemente los parámetros de transmisión de cada canal, por ejemplo variando adaptativamente la modulación o la codificación de los símbolos que se envían en cada uno, en función de las características de cada canal.

En todo caso, y como consecuencia del efecto del canal, los diferentes símbolos recibidos se verán escalados en amplitud y rotados en fase respecto a los símbolos enviados, de modo que será necesario en recepción aplicar mecanismos para compensar dichas variaciones. Una posibilidad, habitualmente empleada, es multiplicar cada uno de los símbolos detectados \tilde{d}_m por un factor que compense la respuesta frecuencial del canal a cada frecuencia. Este proceso se corresponde con una ecualización en el dominio de la frecuencia [3], en tanto que permite llevar a cabo la ecualización del canal no como un filtrado en el dominio temporal, como se realiza típicamente en la mayoría de sistemas de comunicaciones, sino como un proceso efectuado directamente sobre el dominio de la frecuencia, simplemente consistente en la multiplicación de cada símbolo por un coeficiente dependiente de la respuesta frecuencial del canal asociada a dicho símbolo. A modo de ejemplo, en el caso de una ecualización de mínimo error cuadrático medio MMSE (*Minimum Mean Square Error*), el valor del coeficiente por el que se multiplicaría el símbolo \tilde{d}_m en un canal con potencia de ruido P_n sería [2]:

$$W_m = \frac{H^*(m\Delta f)}{|H(m\Delta f)|^2 + P_n} \quad (4.18)$$

En relación a la ecualización en el dominio temporal, la ecualización en el dominio de la frecuencia permite en general una menor complejidad de cálculo, particularmente en canales muy dispersivos como ocurre cuando se pretende efectuar una transmisión de banda ancha. Por este motivo, la técnica de transmisión OFDM resulta eficiente desde la perspectiva de poder utilizar anchos de banda elevados con incremento reducido de la complejidad de los ecualizadores. A modo de ejemplo, en [4] se presenta la comparativa entre la ecualización en el dominio temporal y en el dominio frecuencial en términos del número de multiplicaciones requeridas por símbolo en función del grado de distorsión del canal, llegándose a la conclusión de que la complejidad de la ecualización en el dominio frecuencial se incrementa mucho más lentamente con la longitud de la respuesta impulsional del canal que la ecualización en el dominio temporal, pudiendo llegar a ser la complejidad en esta última hasta un orden de magnitud superior a la de la ecualización en el dominio frecuencial.

Por otra parte, para llevar a cabo la ecualización es necesario disponer en el receptor de mecanismos apropiados de estimación de la respuesta frecuencial del canal. Una forma de llevar a cabo este proceso, tradicionalmente empleada en los sistemas de comunicaciones, consistiría en estimar la respuesta impulsional del canal en el dominio temporal, $h(\tau)$, y tra-

ducirla al dominio de la frecuencia, $H(f)$. Sin embargo, dadas las características de la técnica OFDM y en concreto el comportamiento reflejado en la expresión 4.17, resulta mucho más sencillo efectuar la estimación de los términos $H(m\Delta f)$ directamente en el dominio de la frecuencia, sin pasar por el dominio temporal. Para ello, es suficiente emplear símbolos piloto o de referencia, que son símbolos conocidos a priori por el receptor y que se transmiten en subportadoras e instantes de tiempo específicos. De este modo, de acuerdo con (4.17), el receptor conoce a priori el valor del símbolo enviado $d_m^{(0)}$ que, combinado con el detectado $\tilde{d}_m^{(0)}$, permitiría estimar el valor de $H(m\Delta f)$. Este valor sería extrapolable para instantes de tiempo y frecuencias próximas a las utilizadas por el símbolo de referencia, dependiendo esta "proximidad" de cuan selectivo en frecuencia y tiempo sea el canal. Así, distribuyendo una densidad de símbolos de referencia suficientemente grande que cubrieran diferentes tiempos y frecuencias, se podría obtener la estimación de la respuesta frecuencial del canal para las diferentes subportadoras involucradas.

4.2 OFDMA como técnica de acceso múltiple para el enlace descendente

La técnica de acceso múltiple OFDMA (*Orthogonal Frequency Division Multiple Access*), utilizada en el enlace descendente de LTE, surge de forma natural a partir de la modulación OFDM presentada anteriormente al considerar la posibilidad de que los diferentes símbolos modulados sobre las subportadoras pertenezcan a usuarios distintos. De esta forma, es posible acomodar varias transmisiones simultáneas correspondientes a diferentes flujos de información al viajar en subportadoras diferentes, tal y como se ilustra en el esquema de transmisión de la Figura 4.14. Obsérvese que en dicha figura existen U flujos de información correspondientes a diferentes usuarios, siendo N_k el número de símbolos enviado para el usuario k -ésimo, y donde $d_{i,k}$ representa el i -ésimo símbolo del k -ésimo usuario. Análogamente, en el receptor de cada usuario bastará con recuperar el contenido de las subportadoras asignadas a dicho usuario para separar la información destinada a este usuario de la del resto (obsérvese que será preciso disponer de los mecanismos de señalización adecuados para notificar a cada usuario a través de qué subportadoras se le está enviando la información).

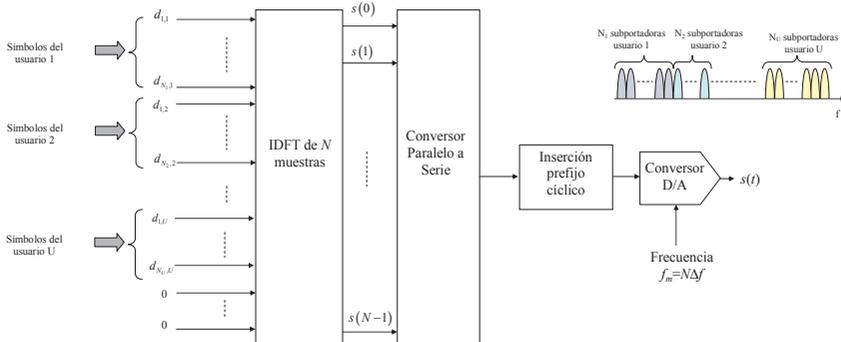


Figura 4.14 Multiplexación de usuarios en OFDMA

Es importante remarcar el hecho de que, si bien en la Figura 4.14 se ilustra el caso de que cada usuario utiliza un conjunto de subportadoras contiguas, esto no tendría por qué ser así, sino que también se podrían distribuir los símbolos de un usuario sobre subportadoras no contiguas.

El empleo de la técnica de acceso múltiple OFDMA conlleva las siguientes ventajas:

- **Diversidad multiusuario:** Mediante OFDMA la asignación de subportadoras a usuarios se lleva a cabo dinámicamente, pudiéndose cambiar en períodos cortos de tiempo dicha asignación a través de estrategias de *scheduling* tal y como se discutirá en el apartado 4.2.1.1. De esta forma, teniendo en cuenta que el canal radio presentará desvanecimientos aleatorios en las diferentes subportadoras, y que dichos desvanecimientos serán independientes para cada usuario, se puede intentar seleccionar para cada subportadora aquél usuario que presente un mejor estado del canal (esto es, perciba una mejor relación señal a ruido), lo que se traducirá en una mejor utilización de la banda disponible para conseguir una mayor velocidad de transmisión, esto es, una mayor eficiencia espectral. Este procedimiento se suele denominar *scheduling* en el dominio de la frecuencia. A modo de ejemplo, en [5] se concluye que gracias a un *scheduling* eficiente de acuerdo con el estado de las diferentes subportadoras, se pueden llegar a obtener ganancias del orden del 40-50% en términos del *throughput* respecto del caso en el que el *scheduling* no explota las variaciones del canal en las subportadoras.
- **Diversidad frecuencial:** Tal y como se ha comentado, es posible asignar a un mismo usuario subportadoras no contiguas, suficientemente separadas como para que el estado del canal en las mismas sea independiente, lo que proporciona diversidad frecuencial en la transmisión de dicho usuario ante canales selectivos en frecuencia.
- **Robustez frente a la propagación multicamino:** Gracias a la aplicación del prefijo cíclico, la técnica OFDMA es muy robusta frente a la interferencia intersimbólica resultante de la propagación multicamino y se puede combatir la distorsión mediante técnicas de ecualización en el dominio de la frecuencia, que resultan más eficientes y menos complejas que las técnicas de ecualización clásicas en el dominio temporal, particularmente cuando se tiene que efectuar una transmisión de banda ancha y en consecuencia se está ante canales muy dispersivos. Esto es particularmente relevante cuando se pretenden emplear bandas de transmisión superiores a 5 MHz, como ocurre con LTE, en que se pretende llegar hasta los 20 MHz. Por ejemplo, en [4] se recogen reducciones de la complejidad computacional de hasta un orden de magnitud en el número de operaciones necesarias para una ecualización en el dominio de la frecuencia en relación a la ecualización en el dominio temporal.
- **Flexibilidad en la banda asignada:** La técnica OFDMA proporciona una forma sencilla de acomodar diferentes velocidades de transmisión a los diferentes usuarios en función de los requerimientos de servicio de cada uno, simplemente a base de la asignación

de más o menos subportadoras por usuario. Nótese que esto puede llevarse a cabo sin ninguna modificación en el proceso de modulación, solamente cambiando los valores de entrada sobre los que se efectúa la IDFT.

- Elevada granularidad en los recursos asignables: Al subdividir la banda total en un conjunto elevado de subportadoras de banda estrecha que se asignan dinámicamente a los usuarios, se dispone de una elevada granularidad a la hora de asignar más o menos recursos a cada uno, lo que resulta útil para acomodar servicios con diferentes requerimientos de calidad. Por el contrario, en otras técnicas como CDMA (*Code Division Multiple Access*), la variación en la cantidad de recursos asignables a cada usuario se consigue principalmente mediante la modificación del factor de ensanchado (*spreading factor*), que acostumbra a tener una granularidad menor en tanto que suele variar en potencias de 2.
- Elevado grado de utilización de la banda asignada: Gracias al empleo de la transmisión OFDM la transmisión multiportadora se consigue con un espaciado mínimo entre las diferentes subportadoras utilizadas, existiendo de hecho una cierta superposición en el espectro ocupado por éstas (ver Figura 4.1) sin que ello afecte a la recuperación de la señal transmitida. Por el contrario, otras técnicas de transmisión diferentes de OFDM requerirían una cierta banda de guarda entre los diferentes canales, lo que reduciría el grado de utilización de la banda asignada y en consecuencia la eficiencia espectral.
- Sencillez de implementación en el dominio digital: Esto es así gracias a la posibilidad de emplear chips que efectúan los procesos de FFT/IFFT de forma rápida.

En todo caso, también es preciso remarcar algunas de las desventajas que esta tecnología presenta:

- Elevada relación entre la potencia instantánea y la potencia media: Tal y como se detallará en el apartado 4.3.1, uno de los inconvenientes de las técnicas de transmisión multiportadora como OFDMA es que la potencia instantánea transmitida puede ser significativamente superior a la potencia media, lo que plantea problemas de linealidad para los amplificadores de potencia, limitando su eficiencia e incrementando su coste.
- Susceptibilidad frente a errores de frecuencia: La ortogonalidad en OFDMA se basa, tal y como se explicó en el apartado 4.1.1, en que la separación entre subportadoras coincide con el inverso de la duración del símbolo OFDMA. Por este motivo, en el caso de que existan desplazamientos en la frecuencia de las subportadoras respecto de su frecuencia de referencia, esto se traducirá en una cierta pérdida de ortogonalidad y la consiguiente interferencia entre subportadoras. Los motivos de estos errores en frecuencia pueden ser diversos, como por ejemplo la estabilidad de los osciladores, el efecto Doppler asociado al movimiento de los terminales, etc. En este sentido, la disponibilidad de mecanismos de sincronización eficientes para hacer frente a estos errores en frecuencia es uno de los aspectos necesarios a tener en cuenta.

4.2.1 Procedimientos de RRM asociados a OFDMA

Para conseguir un uso eficiente de los recursos radio disponibles, que en el caso de OFDMA son las diferentes subportadoras en que se subdivide la banda total, es preciso disponer de un conjunto de mecanismos de gestión para decidir cómo dichos recursos se distribuyen entre los diferentes usuarios y los servicios asociados, según sus requerimientos de calidad de servicio (QoS: *Quality of Service*). Estos mecanismos forman parte de la denominada gestión de recursos radio, en inglés *Radio Resource Management (RRM)*. Si bien en el Capítulo 6 se proporcionará una visión más amplia y detallada de los mecanismos RRM existentes en el contexto de LTE, a continuación se enumeran, con una breve descripción, aquellos mecanismos más íntimamente ligados con las características de la técnica de acceso múltiple OFDMA para el enlace descendente.

4.2.1.1 Scheduling de paquetes

Como se ha comentado, mediante OFDMA es muy sencillo efectuar una asignación dinámica de las subportadoras disponibles a los diferentes usuarios simplemente modificando los símbolos que se inyectan a cada una de las entradas del proceso de IDFT en transmisión. Esto permite que, en periodos muy cortos de tiempo, típicamente compuestos por un pequeño número de periodos de símbolo (e.g., 6 ó 7 periodos en el caso de LTE), se puedan modificar las subportadoras empleadas por cada usuario, lo que proporciona la flexibilidad necesaria para poder acomodar flujos de información con diferentes requerimientos de QoS.

En este contexto, el mecanismo de *scheduling* de paquetes es el responsable de determinar, en cada momento, cuales de las subportadoras se asignan a cada uno de los diferentes usuarios, o incluso a los diferentes flujos de información que pudieran existir de un mismo usuario, correspondientes a diferentes servicios. El proceso se ilustra gráficamente en la Figura 4.15. Como puede apreciarse, por un lado existen datos asociados a cuatro usuarios diferentes, cada uno representado mediante un color, y por el otro el conjunto de recursos se puede contemplar como una rejilla en los ejes frecuencial y temporal, de modo que en frecuencia existe un conjunto de subportadoras separadas Δf , cada una de las cuales se puede asignar durante un período temporal T . El *scheduling* de paquetes es responsable de asociar las parejas subportadora/periodo de tiempo a cada uno de los usuarios.

Para llevar a cabo la decisión de la asignación, en general el algoritmo de *scheduling* tendrá en cuenta información sobre lo siguiente:

- Requisitos de QoS para los diferentes usuarios, en tanto que servicios que requieran de enviar más información deberán disponer de más subportadoras asignadas durante más tiempo.
- Información sobre el estado del canal para cada usuario en las diferentes subportadoras, de modo que se pueda evitar la asignación a un usuario de una subportadora en la que se sabe que el canal presenta una elevada atenuación o interferencia. Notar que

para disponer de esta información, que se puede obtener de medidas efectuadas en el receptor del terminal móvil, será preciso disponer de la señalización apropiada en el enlace ascendente que permita enviar esta información desde el móvil hasta el nodo de la red en el que se ejecute el *scheduling* (y que en LTE será el eNB).

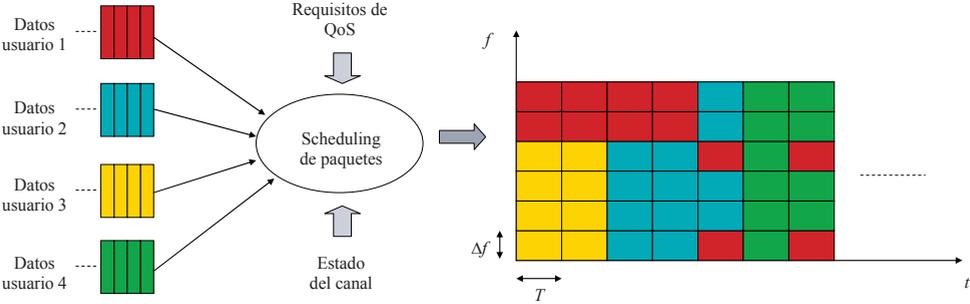


Figura 4.15 Ilustración del scheduling de paquetes en OFDMA

4.2.1.2 Adaptación de enlace

Una de las características de OFDMA es que no impone a priori ninguna condición sobre los símbolos que se modulan sobre las diferentes subportadoras (esto es, los símbolos $d_{i,k}$ de la Figura 4.14), y en consecuencia pueden pertenecer a modulaciones que incluyan más o menos bits de información por símbolo según el orden de la modulación empleada, reflejado en el número de símbolos de su constelación. La Figura 4.16 ilustra a modo de ejemplo las constelaciones para las modulaciones QPSK (*Quadrature Phase Shift Keying*) y 16-QAM (donde QAM responde a las siglas *Quadrature Amplitude Modulation* y donde 16 es el número de símbolos de la constelación asociada a esta modulación). Como puede apreciarse, la constelación define por un lado las amplitudes de la componente en fase (I) y cuadratura (Q) asociadas a cada símbolo o punto, y por otro lado la correspondencia entre símbolos y grupos de bits.

Puede verse como, en el caso de utilizarse una modulación QPSK, cada uno de los símbolos corresponde a 2 bits de información, mientras que en el caso de emplearse una modulación 16-QAM cada símbolo corresponde a 4 bits. Igualmente, en el caso genérico de una modulación m -QAM con m símbolos cada símbolo correspondería a $\log_2 m$ bits.

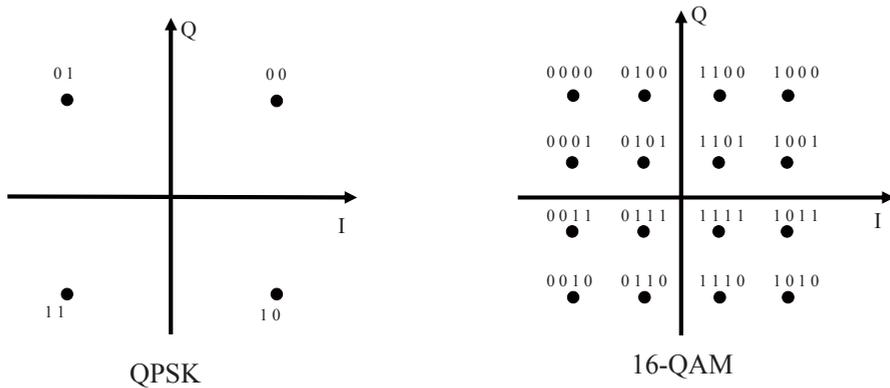


Figura 4.16 Ilustración de las constelaciones para las modulaciones QPSK y 16-QAM

De esta forma, la velocidad de transmisión en bits/s que finalmente resultaría sobre una subportadora con una duración de símbolo T_S y prefijo cíclico T_P y con una modulación que empleara una constelación de m símbolos sería:

$$R(b/s) = \frac{\log_2 m}{T_S + T_P} \quad (4.19)$$

Desde esta perspectiva, con objeto de incrementar la velocidad de transmisión resultaría conveniente incrementar número de bits por símbolo de la modulación empleada a base de utilizar constelaciones con más símbolos. Sin embargo, ante unas condiciones de relación señal a ruido (SNR: *Signal to Noise Ratio*) en el canal dadas, el empleo de modulaciones de orden elevado ocasiona un peor comportamiento en términos de probabilidad de error de bit, ya que al existir más símbolos en la constelación y encontrarse éstos más próximos, es más fácil que debido al ruido se detecte erróneamente un símbolo en lugar de otro. Por este motivo, para poder emplear satisfactoriamente modulaciones con un número elevado de símbolos, es preciso disponer de buenas condiciones de relación señal a ruido. Por lo general, dada una modulación, ésta podrá emplearse apropiadamente con una tasa de error acotada, en tanto que la relación señal a ruido existente en el canal esté por encima de un cierto umbral mínimo, que será mayor cuanto mayor sea el orden de la modulación.

De acuerdo con lo anterior, una estrategia que habitualmente se emplea en el contexto de OFDMA es la denominada adaptación de enlace (o en inglés *link adaptation*), que intenta extraer el máximo rendimiento del canal (en términos de velocidad de transmisión) mediante la selección de la modulación que permita enviar el mayor número de bits por símbolo ante unas condiciones de SNR dadas. De esta forma, aquellas subportadoras que presenten una relación señal a ruido mejor tenderán a utilizar modulaciones de mayor orden que las que presenten una relación señal a ruido menor.

El proceso de adaptación de enlace para OFDMA se ilustra gráficamente en la Figura 4.17,

en un ejemplo en el que existen tres modulaciones disponibles, QPSK, 16-QAM y 64-QAM. La selección de la modulación se efectuaría previo al proceso de la IDFT en función de la SNR medida para las subportadoras asignadas al usuario en cuestión. En el ejemplo ilustrado, se observan a la salida cuatro subportadoras, de las cuales una presenta un nivel muy bajo de SNR, requiriendo por lo tanto una modulación QPSK, otra presenta un nivel elevado, pudiendo emplear una modulación 64-QAM, mientras que las otras dos presentan un nivel intermedio, empleando 16-QAM.

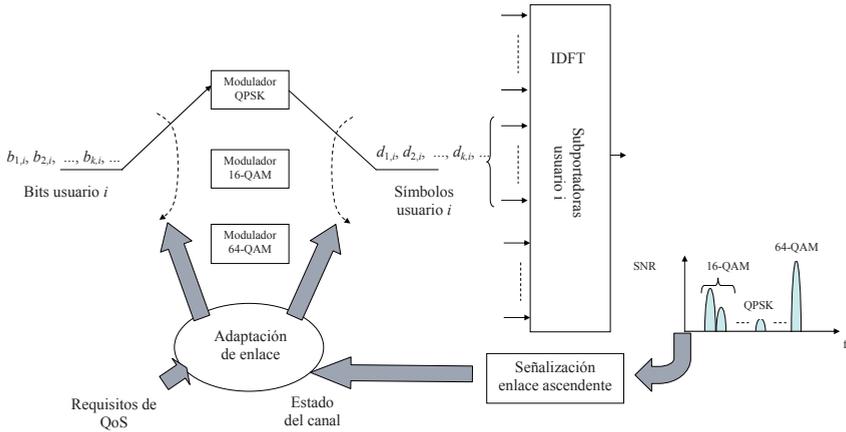


Figura 4.17 Ilustración de la adaptación de enlace en OFDMA

La información que necesita el mecanismo de adaptación de enlace para tomar la decisión apropiada se basa, por un lado en el estado del canal, en términos de SNR, que deberá proporcionarse a través de mecanismos de señalización en el enlace ascendente, pero por otro lado también en los requisitos de QoS, en tanto que por ejemplo determinados servicios con bajos requerimientos de velocidad de transmisión pueden no necesitar de las modulaciones de mayor orden.

Es importante remarcar que el mecanismo de adaptación del enlace debe estar coordinado con el proceso de *scheduling* de paquetes descrito en el apartado anterior, en tanto que este último decide cuales son las subportadoras a asignar a cada uno de los usuarios en cada instante de tiempo.

Por otra parte, también conviene tener en cuenta que, si bien el mecanismo de adaptación de enlace descrito aquí únicamente ha contemplado la selección de la modulación, desde una perspectiva más amplia, también puede seleccionar la codificación de canal empleada, esto es, el grado de redundancia en los diferentes bits enviados para detectar y corregir los posibles errores en la transmisión. Una misma modulación puede combinarse con diferentes códigos de canal y en consecuencia disponer de mayor o menor robustez frente a errores, aunque esto afectará a la velocidad de transmisión neta según la redundancia añadida por el código. Por ejemplo, en el caso de emplearse una modulación con una constelación de m símbolos y una codificación de canal de tasa r , lo que significa que por cada

bit de información útil se están transmitiendo un total de $1/r$ bits, la velocidad neta que se conseguiría es:

$$R(b/s) = \frac{r \log_2 m}{T_s + T_p} \quad (4.20)$$

Claramente, cuanto menor sea la tasa del código r (es decir, hay una mayor redundancia) mayor será su robustez para corregir errores (y por lo tanto podrá trabajar en condiciones de SNR más reducida), pero esto reducirá la velocidad neta que se puede conseguir. De acuerdo con esto, desde una perspectiva genérica, dadas unas condiciones de SNR, el mecanismo de adaptación de enlace intentará escoger la combinación de codificación de canal y modulación (denominada en inglés MCS: *Modulation and Coding Scheme*) que permita conseguir la máxima velocidad de transmisión neta.

4.2.2 Sistemas celulares OFDMA

El despliegue de un sistema de comunicaciones móviles para proporcionar cobertura a una determinada región se basa en la distribución de un conjunto de estaciones base o células por el territorio a cubrir, de modo que cada una proporciona la cobertura a los usuarios de una zona geográfica, y entre todas las bases aseguran la cobertura de todo el territorio. Esto da lugar a lo que tradicionalmente se ha denominado un sistema celular, y que se ha venido empleando en todos los sistemas de comunicaciones móviles previos a LTE, tales como GSM o UMTS.

Uno de los problemas a los que se debe hacer frente en un sistema celular consiste en determinar qué recursos radio se asignan a las diferentes estaciones base para conseguir, por un lado, disponer de la suficiente capacidad de acuerdo con la cantidad de usuarios a los que cada base debe dar servicio, y por el otro lado conseguir una utilización eficiente de dichos recursos, intentando consumir el mínimo número de ellos.

Esta asignación de recursos a bases está completamente ligada a la técnica de acceso múltiple empleada. Así, por ejemplo en el caso de GSM, que emplea una combinación de acceso múltiple por división en tiempo y en frecuencia (TDMA/FDMA), se suele asignar a cada estación base un conjunto de radiocanales diferente que pueden reutilizarse en otras estaciones base que se encuentren a una cierta distancia de la base en cuestión, para asegurar que no haya interferencia intercelular apreciable. Por el contrario, en el caso de UMTS, que emplea la técnica de acceso múltiple por división en código (CDMA), todas las bases utilizan la misma frecuencia pero con diferentes secuencias código (los denominados códigos de *scrambling*). En este caso, además de la diferenciación por código, con objeto de acotar la interferencia intercelular, es preciso el empleo de técnicas de control de la potencia transmitida.

Por su parte, en el caso de los sistemas celulares basados en OFDMA, en los que la banda total se subdivide en un conjunto de subportadoras, y por lo tanto presentan intrínsecamente

una componente FDMA, una forma natural de proceder consistiría en emplear técnicas de reutilización de frecuencia similares a las de los sistemas de segunda generación como GSM. En este caso, el conjunto total de subportadoras disponible se subdivide en F grupos (siendo F el factor de reuso) cada uno de los cuales se asigna a una célula diferente de un conjunto de F células denominado *cluster*. A su vez, cada uno de los grupos se reutiliza en células pertenecientes a diferentes *clústeres*. Este proceso se ilustra en la Figura 4.18 para un caso de factor de reuso $F=3$ asumiendo que cada base se encarga de dar servicio a un área hexagonal. Como puede apreciarse, mediante este mecanismo se garantiza que células adyacentes no trabajen con las mismas subportadoras y por lo tanto no se interfieran mutuamente.

Obsérvese que, a base de incrementar el valor de F , se consigue alejar más las células que utilizan la misma frecuencia y por lo tanto reducir la interferencia intercelular. Sin embargo, esto es a costa de que el número de subportadoras por célula N/F se reduce, lo cual se traducirá en una reducción de la capacidad de la célula, por lo que en general no interesará utilizar valores de F muy elevados.

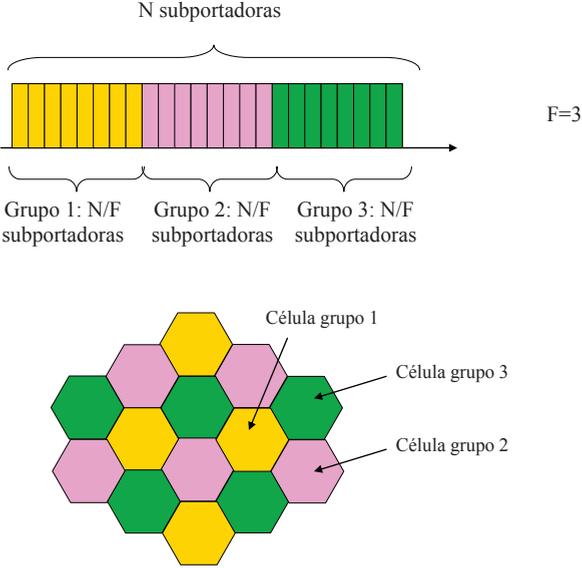


Figura 4.18 Ejemplo de asignación de subportadoras a células según un factor de reuso $F=3$

Por otra parte, es importante remarcar que el empleo de las técnicas de *scheduling* de paquetes y de adaptación de enlace descritas en los apartados anteriores permite efectuar dentro de cada célula una gestión dinámica de las subportadoras disponibles, asignándolas a los usuarios según sus condiciones de canal medidas ahora no sólo como relación señal a ruido sino como relación señal a ruido más interferencia. De esta forma, a diferencia de lo que ocurría en los sistemas de segunda generación como GSM, en OFDMA la interferencia intercelular no se controla únicamente mediante el reuso de frecuencias sino que también se puede regular mediante un *scheduling* apropiado. Por este motivo, sería posible trabajar con factores

de reuso inferiores, pudiendo llegar incluso a utilizarse un factor $F=1$, esto es, asignar todas subportadoras disponibles en todas las células. En esta situación, sería tarea del *scheduling* asignar apropiadamente las subportadoras a los usuarios, de modo que a aquellos usuarios más susceptibles a la interferencia intercelular, típicamente aquellos que se encuentren en el extremo de la célula, se les asignasen subportadoras que en un momento dado no están siendo utilizadas en las células adyacentes, mientras que a los usuarios menos susceptibles a la interferencia intercelular, típicamente los que se encuentran cerca de la estación base, se les podrían asignar subportadoras utilizadas en otras células.

La operación de un sistema celular OFDMA con reuso $F=1$ se ilustra en la Figura 4.19. Como puede observarse, los terminales UE1 y UE2 se encuentran conectados a la célula 1 y el terminal UE3 a la célula 2. Los terminales UE2 y UE3 se encuentran muy cerca de sus respectivas estaciones base, de modo que ambas estaciones base pueden utilizar el mismo conjunto de subportadoras para ambos usuarios. Por el contrario, para el caso del terminal UE1, que se encuentra ubicado en el extremo de la célula 1, el *scheduling* deberá seleccionar un conjunto de subportadoras que no estén siendo utilizadas por la célula 2, ya que en caso contrario, el terminal UE1 podría percibir un exceso de interferencia.

Claramente, el empleo de un reuso completo $F=1$ permitiría a una célula disponer potencialmente del máximo de capacidad de transmisión, en tanto que podría llegar a asignar todas las subportadoras disponibles a un único usuario que se encontrara cerca de la estación base y que por lo tanto podría obtener una velocidad de transmisión superior al caso de utilizarse $F>1$. Sin embargo, también es cierto que este reuso completo en general tendería a penalizar a los usuarios en el extremo de la célula, al margen de requerir mecanismos de *scheduling* y de coordinación entre diferentes bases más sofisticados que si se utilizara por ejemplo $F>1$. Dentro del Capítulo 6 se presentarán, en el contexto de LTE, estrategias de reuso de frecuencias más sofisticadas que el simple reuso fijo con un factor F , a la vez que se discutirán los aspectos relacionados con las técnicas de coordinación de la interferencia intercelular.

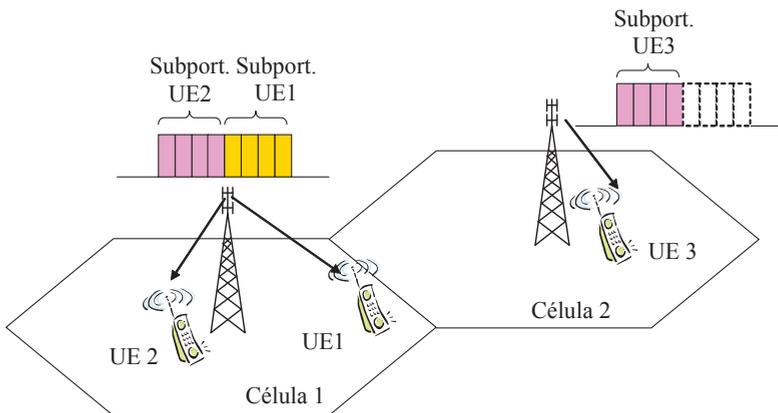


Figura 4.19 Ejemplo de OFDMA con reuso $F=1$

Un caso particular de redes celulares OFDMA son las denominadas redes isofrecuenciales, en inglés *Single Frequency Networks* (SFN), que pueden utilizarse para sistemas de transmisión *broadcast o multicast*, es decir para entregar la misma información a diferentes usuarios que se ubican en células diferentes, por ejemplo el caso de servicios de televisión, servicios de difusión de información como noticias, etc.

En este tipo de sistemas, todas las células involucradas transmiten la misma información empleando las mismas subportadoras, tal y como se ilustra en la Figura 4.20, en la que un móvil recibe la señal proveniente de dos estaciones base. En esta situación, la señal recibida en el móvil será la suma de varias réplicas de la misma señal pero retardadas y con diferentes amplitudes y fases, según la trayectoria recorrida durante la propagación de la señal, tal y como también se ilustra en la figura. De este modo, desde la perspectiva del móvil, esto sería equivalente al caso en que la transmisión se hubiera efectuado desde una única base pero estuviera sujeta a la propagación multicamino, de acuerdo con una cierta respuesta impulsional. Dicho de otro modo, para el terminal móvil sería imposible conocer si las réplicas recibidas corresponden a la señal transmitida por otra base o bien si son fruto de la propagación multicamino de únicamente su propia base. Por este motivo, gracias al empleo del prefijo cíclico, el terminal móvil será capaz de recuperar automáticamente la señal originalmente transmitida, tal y como se explicó en el apartado 4.1.2. Naturalmente, para poder emplear satisfactoriamente esta técnica, es preciso que la duración del prefijo cíclico sea suficientemente larga como para cubrir las diferencias en los tiempos de propagación desde las distintas bases, lo que en general requerirá de prefijos cíclicos más largos que en las redes celulares OFDMA convencionales.

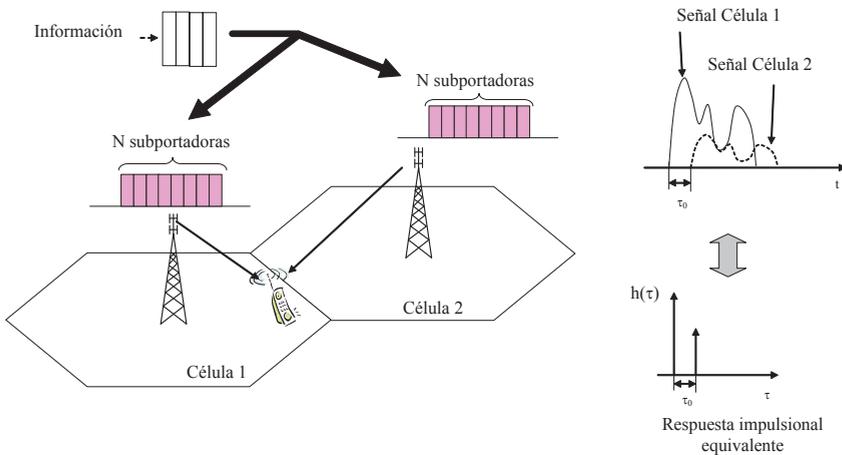


Figura 4.20 Sistema SFN

4.2.3 Parámetros de OFDMA empleados por LTE

Con objeto de ilustrar los conceptos anteriormente presentados para la técnica de acceso

múltiple OFDMA, a continuación se presentan los parámetros definidos en el contexto del enlace descendente del sistema LTE.

La separación definida entre subportadoras es de $\Delta f=15$ kHz, si bien también existe un espaciado reducido de 7.5 kHz, particularmente pensado para el caso de transmisión multicast en SFN.

Las subportadoras se agrupan en bloques de 12 subportadoras consecutivas, con un ancho de banda resultante de $12\Delta f=180$ kHz, constituyendo cada uno de estos bloques un recurso en el dominio frecuencial, o equivalentemente la unidad mínima de asignación de recursos a un usuario. A partir de aquí, el número total de subportadoras ocupadas por una portadora LTE en el enlace descendente es de $N_s=12N_b+1$, siendo N_b el número de bloques de 12 subportadoras utilizados. En esta última expresión, hay que tener en cuenta que la subportadora en el centro de la banda no se utiliza en ninguno de los bloques, en tanto que puede ser altamente interferida por el oscilador local en recepción.

La flexibilidad en el ancho de banda ocupado viene dada por el número de bloques N_b utilizados, que puede tomar los valores {6, 15, 25, 50, 75, 100}, correspondiente a anchuras de banda de transmisión de $BW=(12N_b+1)\Delta f=\{1.095, 2.715, 4.515, 9.015, 13.515, 18.015\}$ MHz. De acuerdo con estos valores, se pueden configurar diferentes espaciados entre canales LTE, correspondientes a {1.4, 3, 5, 10, 15, 20} MHz, en los que, como puede observarse, la anchura de banda ocupada es aproximadamente un 90% del espaciado entre canales (excepto para el caso de 1.4 MHz, que es del 78%).

El empleo de una u otra configuración en términos del número de bloques empleados se regula mediante el número de muestras empleado para los procesos de IDFT/DFT, dados por: $N=\{128, 256, 512, 1024, 1536, 2048\}$, y los correspondientes valores de frecuencia de muestreo $f_m=N\Delta f=\{1.92, 3.84, 7.68, 15.36, 23.04, 30.72\}$ MHz. Es importante remarcar que la selección de estos valores de frecuencia de muestreo se hizo teniendo en cuenta aspectos de compatibilidad con los actuales sistemas UMTS, en tanto que la frecuencia de muestreo de 3.84 MHz coincide con la tasa de chips de UMTS de valor 3.84 Mchips/s, lo que facilita la implementación de terminales multimodo UMTS/LTE con un único oscilador.

Por otra parte, del conjunto anterior de frecuencias de muestreo, la mayor de ellas define el intervalo temporal de referencia como $T_m=1/30.72$ μ s, a partir del cual se definen las duraciones de los símbolos, de los prefijos cíclicos, y finalmente de la estructura temporal de trama que se detallará en el Capítulo 5.

En concreto, la duración del símbolo OFDM, que corresponde al inverso de la separación entre subportadoras, $T_s=1/\Delta f=66.67$ μ s, puede expresarse como $T_s=2048 T_m$.

Análogamente, para la duración del prefijo cíclico existen diferentes valores, que son $T_p=160T_m=5.21$ μ s, $T_p=144T_m=4.68$ μ s y $T_p=512T_m=16.7$ μ s. La utilización de uno u otro valor se detallará más adelante en el Capítulo 5 cuando se presente el esquema detallado de la capa física en LTE.

Por último, comentar que en relación a las modulaciones empleadas, en el enlace descendente de LTE pueden utilizarse QPSK, 16QAM y 64 QAM, correspondientes a 2, 4 y 6 bits por símbolo, respectivamente.

4.3 SC-FDMA como técnica de acceso múltiple para el enlace ascendente

4.3.1 Motivación

Uno de los aspectos importantes a tener en cuenta en la elección de una modulación para un sistema de comunicaciones es la variación en la potencia instantánea de la señal transmitida, en tanto que si es muy grande la eficiencia de los amplificadores de potencia se reduce. Una métrica para caracterizar esta variación es el denominado factor de cresta o PAPR (del inglés *Peak-to-Average-Power Ratio*), que en definitiva mide la relación entre la potencia instantánea de la señal transmitida respecto de la potencia media, y que se define matemáticamente como [1]:

$$PAPR = \frac{|s(t)|^2}{E[|s(t)|^2]} \tag{4.21}$$

Señales con un valor de PAPR muy grande requieren amplificadores de potencia altamente lineales para evitar la distorsión asociada a la intermodulación. Para conseguir esta elevada linealidad, el amplificador debe operar bastante por debajo de su potencia de pico, lo que se traduce finalmente en una eficiencia reducida, entendida dicha eficiencia como la relación entre la potencia transmitida respecto de la potencia consumida.

En el caso de modulaciones multi-portadora, como ocurre con OFDM (y por extensión con OFDMA), el PAPR se incrementa al incrementar el número de subportadoras utilizadas. Este efecto se ilustra gráficamente en la Figura 4.21 y en la Figura 4.22, que muestran, respectivamente, la evolución temporal de un símbolo OFDM y del PAPR resultante cuando se emplean 3, 6 y 12 subportadoras, moduladas por una secuencia de símbolos alternados +1 y -1. Como puede apreciarse, el incremento en el número de subportadoras se traduce en un incremento en el valor máximo del PAPR que puede llegar a existir.

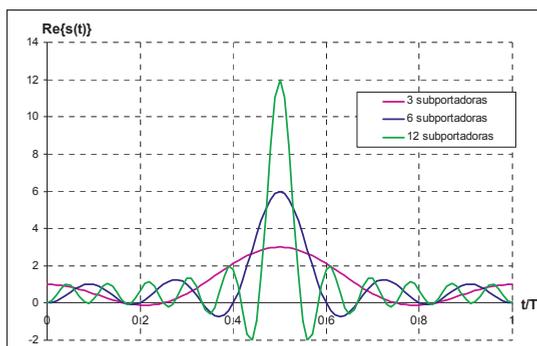


Figura 4.21 Ejemplo de la evolución temporal de un símbolo OFDM compuesto por 3, 6 y 12 subportadoras, moduladas por una secuencia de símbolos alternados +1, -1

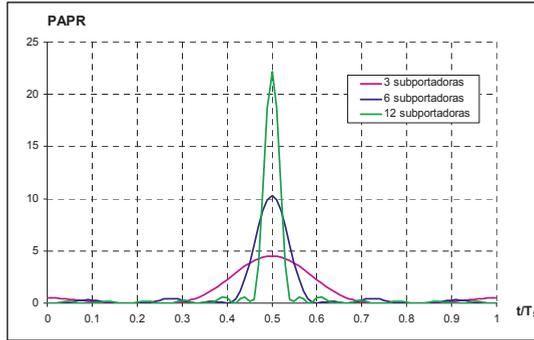


Figura 4.22 Ejemplo del PAPR para un símbolo OFDM compuesto por 3, 6 y 12 subportadoras, moduladas por una secuencia de símbolos alternados +1, -1

En general, si el número de subportadoras es muy grande, como ocurre de hecho en la práctica, esto tenderá a incrementar el PAPR de la señal transmitida. A modo de ejemplo, y de acuerdo con resultados obtenidos en [6], en la Figura 4.23 se muestra la distribución estadística del PAPR en una transmisión OFDMA en términos de la función de distribución acumulada complementaria (CCDF: *Complementary Cumulative Distribution Function*), que refleja para un cierto valor x la probabilidad de que el PAPR sea superior a dicho valor x . La gráfica considera una señal OFDMA construida con una IDFT/DFT de longitud $N=256$ y en la que la transmisión ocupa $K=64$ subportadoras. Como puede apreciarse, existe una probabilidad del 10% de tener un PAPR superior a 8.5 dB, y del 1% de tener un PAPR superior a 10 dB.

También es importante remarcar que, en el caso de OFDMA, el PAPR no es muy sensible al hecho de que los símbolos enviados correspondan a una modulación u otra, a diferencia de los esquemas de transmisión que no emplean múltiples portadoras, en los que por ejemplo el empleo de 16-QAM ocasiona un PAPR superior al empleo de QPSK.

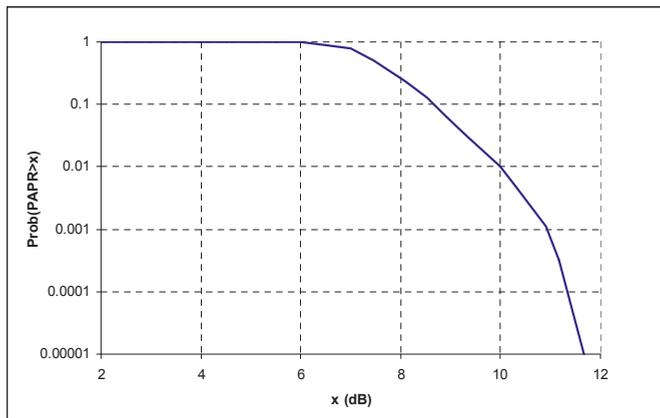


Figura 4.23 CCDF del PAPR para una transmisión OFDMA

Conseguir una eficiencia elevada es algo crítico para los amplificadores de los terminales

móviles, en tanto que esto permite reducir tanto el consumo de potencia del terminal (y consecuentemente incrementar la duración de las baterías) como su coste. Si bien existen algunos métodos que permiten reducir el PAPR de la señal OFDMA (por ejemplo mediante la reserva de algunas subportadoras no utilizadas, o bien mediante técnicas de prefiltrado o precodificación de los símbolos antes de modularlos según OFDM), estos métodos también acostumbran a comportar un incremento significativo en la complejidad computacional o una reducción en las prestaciones obtenidas [2]. Por este motivo, en el sistema LTE, la técnica de acceso múltiple OFDMA se emplea únicamente en el enlace descendente, en tanto que en la estación de base no son tan críticos ni la eficiencia ni el coste de los amplificadores. Por el contrario, en el enlace ascendente, se ha optado por utilizar una técnica de acceso de portadora única (*single carrier*), la denominada SC-FDMA (*Single Carrier Frequency Division Multiple Access*), y que pasamos a describir a continuación.

4.3.2 Esquemas de transmisión y recepción SC-FDMA

A la hora de seleccionar el esquema apropiado para el enlace ascendente de LTE, se intentaron conjugar las siguientes propiedades deseables:

- Variaciones reducidas en la potencia instantánea de la señal transmitida, en términos de un valor de PAPR reducido, para conseguir una mayor eficiencia en los amplificadores de potencia y un menor coste de los mismos.
- Posibilidad de llevar a cabo de forma sencilla mecanismos de ecualización en el dominio de la frecuencia, que, tal y como se ha discutido en el apartado 4.1.2, presentan una menor complejidad que la ecualización en el dominio temporal.
- Capacidad de proporcionar una asignación de banda flexible, que pueda variarse de forma sencilla de acuerdo con las diferentes necesidades de transmisión de los diferentes usuarios.

De acuerdo con los tres requerimientos anteriores, la técnica de acceso múltiple seleccionada SC-FDMA de hecho se basa en unos principios de transmisión muy similares a los de OFDM, pero efectuando una precodificación de los símbolos a transmitir previa al proceso de transmisión OFDM, lo que permite reducir las variaciones en la potencia instantánea. Por otro lado, la capacidad de poder efectuar ecualización en el dominio de la frecuencia es algo inherente a la técnica OFDM con empleo de prefijo cíclico, tal y como se discutió en el apartado 4.1.2, mientras que la flexibilidad en la banda asignada también es intrínseca a OFDM simplemente a base de dedicar más o menos subportadoras a un usuario.

El esquema de transmisión en banda base resultante para SC-FDMA se muestra en la Figura 4.24. Como puede apreciarse, existen K símbolos a transmitir, los cuales son precodificados mediante una DFT de K muestras, como paso previo a efectuar una transmisión OFDM de acuerdo con una IDFT de N muestras, con una separación entre subportadoras Δf , y con el consiguiente añadido del prefijo cíclico. Debido al proceso de precodificación basado en DFT,

esta técnica de transmisión de señal también suele denominarse como *DFT-Spread OFDM* (DFTS-OFDM).

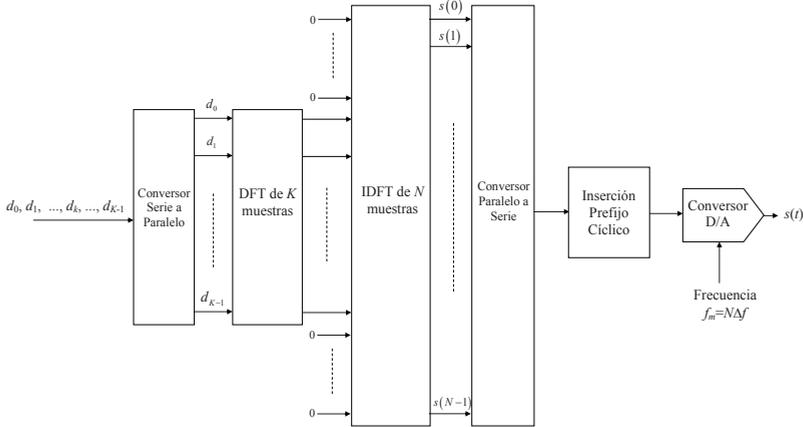


Figura 4.24 Esquema de transmisión de la señal SC-FDMA

Claramente, en el esquema mostrado, en el caso de que el tamaño de la DFT, K , fuera igual al de la IDFT, N , los procesos de DFT y de IDFT se cancelarían entre sí sin tener ningún efecto, por lo que de hecho la señal enviada sería simplemente el mismo conjunto de símbolos original, resultando en una señal en banda base no modulada sobre diferentes subportadoras, esto es, una señal de portadora única (*single carrier*), que presenta mejores propiedades de PAPR que las señales multiportadora. Sin embargo, en tanto que $K < N$ y el resto de entradas de la IDFT están puestas a 0, el resultado del proceso es una señal que continua teniendo la propiedad de ser de portadora única, y cuyo ancho de banda $B = K\Delta f = Kf_m/N$ es regulable simplemente cambiando el valor de K , lo que permite proporcionar la deseada flexibilidad en la banda asignada. Por otra parte, simplemente escogiendo adecuadamente las entradas de la IDFT sobre las que se introducen los K símbolos de salida de la IDFT, se puede cambiar el rango de frecuencias de salida sobre las que se ubicará la señal resultante.

Tal y como ocurría en el caso de OFDMA, la selección del valor de muestras de la IDFT suele tomarse de modo que N sea una potencia de 2, lo que permite implementar la IDFT de forma rápida mediante un proceso de transformada rápida de Fourier inversa (IFFT). A su vez, en el caso de K , si bien también sería deseable esta misma propiedad, resultaría en una reducción del grado de flexibilidad a la hora de asignar la banda de señal (en tanto que solamente se podrían tener como bandas múltiplos de Δf que fueran potencia de 2). Por este motivo, en la práctica se ha optado por permitir que K no sea potencia de 2, y esta condición se sustituye por el hecho de que K se pueda expresar como el producto de números primos pequeños, lo que todavía permite una implementación relativamente rápida del proceso de IDFT [2].

La Figura 4.25 presenta otro esquema posible de transmisión SC-FDMA, consistente en no ubicar las K muestras de salida de la DFT de forma contigua sino en distribuir las K muestras de salida de la DFT de forma equidistante en las entradas de la IDFT, ubicando ceros en las posiciones intermedias. Por este motivo, este

esquema se denomina SC-FDMA distribuido, a diferencia del presentado anteriormente en la Figura 4.24, que se conoce como localizado. Por simplicidad, en la Figura 4.25 se ha obviado la representación explícita de los conversores serie a paralelo y paralelo a serie que se muestran en las otras figuras. Hay que hacer notar que, si bien el esquema distribuido resulta en que la transmisión ocupa diferentes porciones de espectro no contiguas, las variaciones de potencia en términos de PAPR son similares a las del esquema localizado, y por otra parte, el esquema distribuido proporciona una mayor diversidad frecuencial en tanto que la señal se distribuye entre portadoras separadas. Si bien ambos esquemas son viables técnicamente, desde la perspectiva de LTE para el enlace ascendente, ha sido el esquema localizado el seleccionado, en tanto que es menos sensible a los errores de frecuencia.

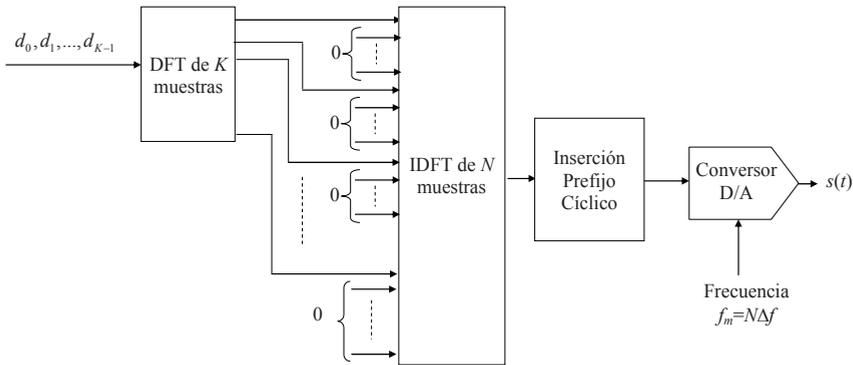


Figura 4.25 Esquema de transmisión SC-FDMA distribuido

El mecanismo de multiplexación de transmisiones de diferentes usuarios según SC-FDMA para el enlace ascendente se ilustra gráficamente en la Figura 4.26, en la que se aprecian las transmisiones de dos usuarios diferentes. Obsérvese como, manteniendo los mismos parámetros a nivel de número de muestras de la IDFT, N , frecuencia de muestreo y separación entre subportadoras Δf , las transmisiones de los dos usuarios vienen ubicadas en diferentes entradas de la IDFT, de modo que en las posiciones de entrada en las que se ubica la transmisión del usuario 2 el usuario 1 inyectará ceros (y a la inversa, el usuario 2 inyectará ceros en las posiciones de entrada en las que se encuentra la señal del usuario 1). Como resultado se tienen dos transmisiones que ocupan bandas frecuenciales diferentes. Obsérvese también como el diferente ancho de banda asignado a los dos usuarios se consigue mediante DFTs de diferente tamaño, una de K_1 muestras para el usuario 1, resultando en una banda de valor $K_1\Delta f$, y otra de K_2 para el usuario 2, resultando en una banda de $K_2\Delta f$.

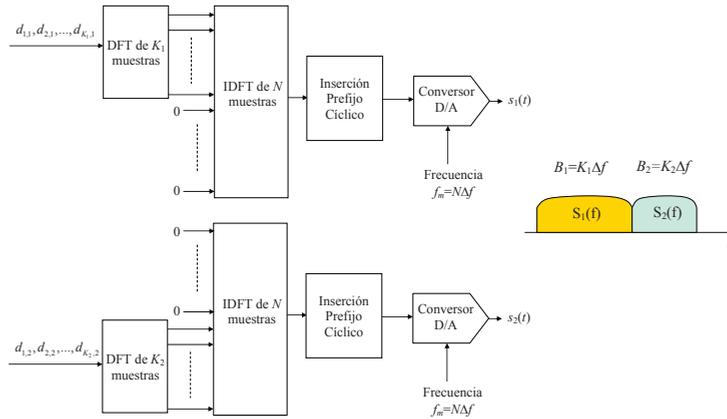


Figura 4.26 Multiplexación de usuarios con SC-FDMA

El esquema de recepción de la señal SC-FDMA se muestra en la Figura 4.27. Tras el proceso de conversión A/D y la eliminación del prefijo cíclico, el proceso de recepción se efectúa siguiendo la misma metodología que en el caso de OFDMA, esto es, mediante una DFT de N muestras. De entre estas muestras se seleccionarán a la salida únicamente las asociadas a las K subportadoras en las que se ubicó la señal en el transmisor (ver Figura 4.24), mientras que el resto de muestras son descartadas en tanto que estarían asociadas a los ceros que se añadieron en el transmisor. El proceso finaliza mediante la realización de la IDFT sobre las K muestras seleccionadas, y que invierte el proceso de precodificación mediante DFT que se realizó en transmisión. De este modo, a la salida se obtiene la estimación de los símbolos enviados.

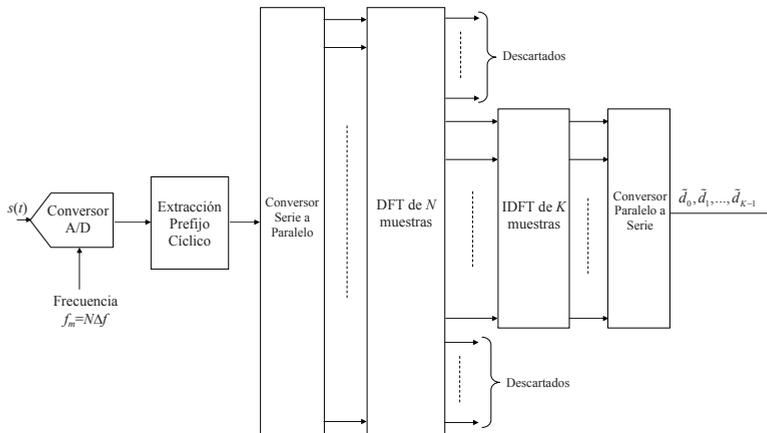


Figura 4.27 Esquema de recepción de la señal SC-FDMA

En el caso de existir canales dispersivos es preciso incorporar en el proceso de recepción un mecanismo de ecualización del canal, con objeto de compensar la distorsión introducida

por el mismo. En esta situación, el diagrama de bloques del receptor es el que se presenta en la Figura 4.28. Gracias al empleo del prefijo cíclico, tal y como se discutió en el apartado 4.1.2, a la salida de la DFT de N muestras la señal en cada una de las K muestras seleccionadas (cada una asociada con una subportadora) se corresponderá con la representación frecuencial de la señal transmitida (esto es, la DFT de los K símbolos que se generó en el transmisor) multiplicada por la respuesta frecuencial del canal a la frecuencia de dicha subportadora, según la ecuación (4.17). De esta forma, la ecualización se puede efectuar directamente en el dominio de la frecuencia, simplemente multiplicando cada una de las K muestras de salida de la DFT por un coeficiente que compense la respuesta frecuencial del canal asociada a dicha muestra o subportadora, por ejemplo según la ecuación (4.18) en el caso de una ecualización de mínimo error cuadrático medio. De este modo, la complejidad computacional del proceso de ecualización es inferior a la que resultaría en caso de tener que implementarlo mediante un filtrado temporal.

Tras el proceso de ecualización, la estimación de los símbolos enviados se obtiene simplemente mediante el proceso de IDFT de K muestras para traducir la señal del dominio de la frecuencia al dominio temporal, tal y como se observa en la Figura 4.28.

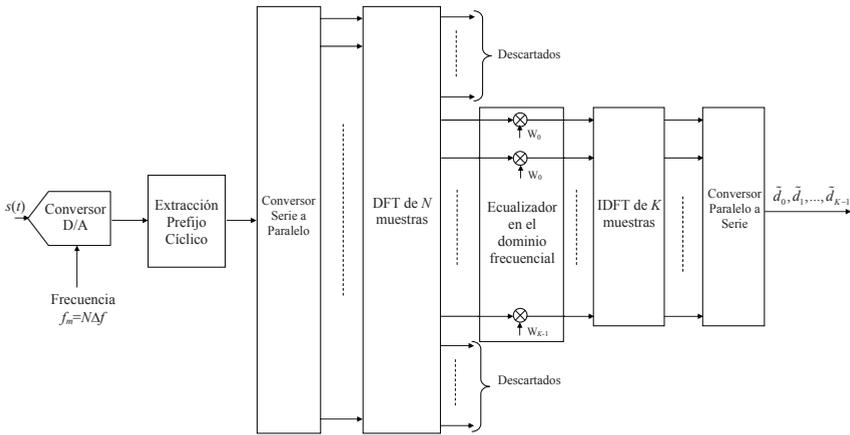


Figura 4.28 Esquema de recepción de la señal SC-FDMA con ecualización en el dominio de la frecuencia

La señal SC-FDMA presenta unas propiedades de PAPR mejores que las de la señal OFDMA, gracias a su característica de transmisión en portadora única. Sin embargo, y a diferencia de lo que ocurre con OFDMA, en que la PAPR es bastante insensible a la modulación empleada, la PAPR de la señal SC-FDMA se degrada más significativamente cuando los símbolos enviados corresponden a modulaciones de órdenes superiores, como 16-QAM y 64-QAM. Esto se ilustra gráficamente en la Figura 4.29, que muestra la distribución estadística del PAPR en términos de la CCDF para los casos de OFDMA y de SC-FDMA con modulaciones QPSK y 16-QAM. Los resultados, extraídos de [6], se han obtenido mediante simulaciones con $N=256$ subpor-

tadoras y $K=64$ subportadoras por usuario. Como puede apreciarse, existe una reducción en el PAPR respecto del caso de emplear OFDMA y esta reducción es sensible a la modulación empleada, siendo mayor en el caso de QPSK.

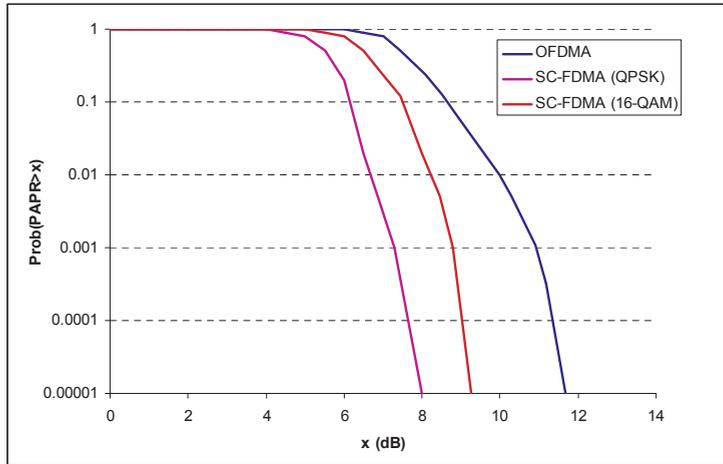


Figura 4.29 Ejemplo de la CCDF del PAPR para OFDMA y para SC-FDMA

Una de las posibilidades empleadas para reducir todavía más la PAPR de la señal SC-FDMA consiste en emplear en transmisión técnicas de conformación de espectro (*spectrum shaping*) [7]. En la Figura 4.30 se muestra el correspondiente esquema de transmisión. Como puede verse, tras efectuar la DFT de K muestras, se aplica el proceso de conformación sobre la señal resultante, previo a efectuar la IDFT. El proceso de conformación consta de dos fases, una primera en la que se efectúa una cierta expansión del espectro original, y otra en la que cada una de las muestras se multiplica por una función de conformación, como por ejemplo la muy habitualmente empleada función coseno realzado. Es importante destacar que, si bien este procesado es capaz de mejorar las prestaciones en términos de variación de potencia instantánea, dicha mejora es a costa de un cierto empeoramiento desde la perspectiva de eficiencia espectral, en tanto que la conformación espectral ocasiona un incremento en el ancho de banda ocupado por la señal transmitida, de acuerdo al denominado factor de *roll-off* de la función de conformación empleada (por ejemplo, un factor de *roll-off* de 0.22 se traduce directamente en un incremento en el 22% en la banda utilizada). En la figura, este incremento de ancho de banda ocupado se ilustra por el hecho de que a la salida del bloque conformador de espectro existe un total de P muestras, superior a las K obtenidas de la DFT.

Por otro lado, otro de los aspectos importantes a tener en cuenta en el enlace ascendente es el control temporal de las señales provenientes de los diferentes móviles. En concreto, y con objeto de preservar la ortogonalidad de las subportadoras empleadas por diferentes terminales, es preciso que los desajustes temporales entre señales sean inferiores a la duración del prefijo cíclico. En caso contrario, esto podría dar lugar a interferencia entre transmisiones.

Este control se lleva a cabo mediante mecanismos de avance temporal (*time advance*), basados en adelantar el instante de transmisión de cada terminal de acuerdo con su retardo de propagación para asegurar que todas las transmisiones lleguen aproximadamente en el mismo instante. Estos mecanismos se verán más en detalle en el Capítulo 5 para el caso de LTE.

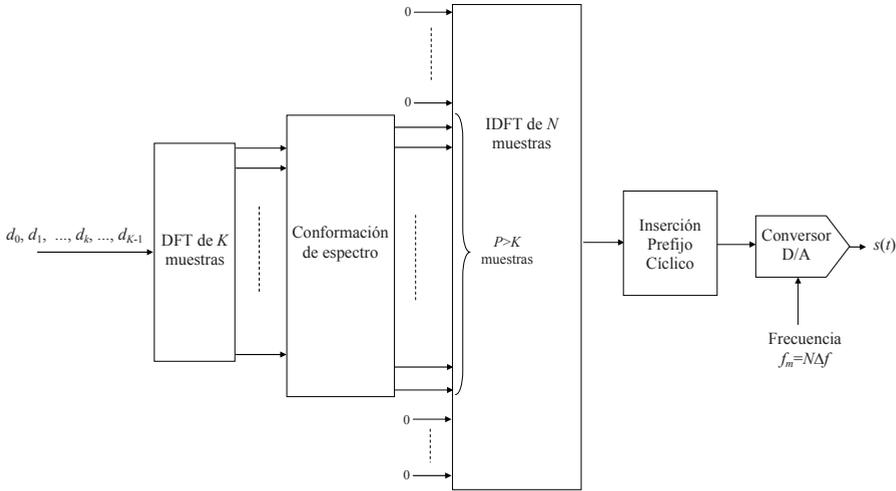


Figura 4.30 Esquema de transmisión de la señal SC-FDMA con conformación de espectro

4.3.3 Procedimientos de RRM asociados a SC-FDMA

Los procedimientos de gestión de recursos asociados al enlace ascendente para SC-FDMA son de hecho similares a los empleados por OFDMA, en tanto que ambas técnicas de acceso se basan en unos principios de transmisión parecidos. En este sentido será preciso, tal y como se detalló en el apartado 4.2.1, disponer también de mecanismos de *scheduling* y de adaptación de enlace para saber cuales son los recursos asignados en cada momento a un usuario, así como los formatos de modulación y codificación de canal a emplear. En todo caso, es preciso destacar que aspectos como la estimación de canal en el enlace ascendente, necesarios para llevar a cabo un *scheduling* que tenga en cuenta las condiciones de canal de cada usuario, son más complejos que en el enlace descendente, ya que implican la necesidad de disponer de señales de referencia que se envíen desde cada uno de los móviles involucrados en el proceso de *scheduling* (a diferencia del enlace descendente, en que los mismos símbolos de referencia podían ser empleados por todos los terminales), lo que se traducirá en un incremento en la señalización del sistema.

4.3.4 Parámetros de SC-FDMA empleados por LTE

Los parámetros definidos para SC-FDMA en el contexto del enlace ascendente para LTE

presentan bastantes similitudes con los del enlace descendente, en tanto que la técnica de transmisión se basa en principios similares. En particular, la separación entre subportadoras es también de $\Delta f=15$ kHz agrupadas en bloques de 12 subportadoras cada uno. Dichos bloques se pueden asignar de forma flexible a los diferentes usuarios según sus necesidades, con la única limitación de que, al emplearse SC-FDMA localizado, todas las subportadoras asignadas a un único usuario deben ser contiguas.

A diferencia del enlace descendente, el número de subportadoras total disponible es de $N_s=12N_B$, ya que ahora no se deja sin utilizar la subportadora central de la banda. Puesto que con SC-FDMA un usuario debe ocupar un conjunto de subportadoras contiguas, gracias a permitir el uso de la subportadora central, sería potencialmente posible llegar a asignar todas las subportadoras de la banda a un único usuario. Por otra parte, la posible interferencia del oscilador local sobre la subportadora central, que era el problema del enlace descendente, gracias a la precodificación basada en la DFT empleada por SC-FDMA quedaría dispersada sobre toda la banda ocupada, por lo que no resulta ser tan crítico como en el enlace descendente.

En relación a los valores del período de símbolo y la duración del prefijo cíclico, son los mismos que se emplean en el enlace descendente.

4.4 Estructuras con varias antenas

La diversidad en espacio es la primera estructura con varias antenas ampliamente usada en comunicaciones móviles [8]. Se basa en el hecho de que un canal móvil se caracteriza por una respuesta impulsional aleatoria y variante no sólo en el dominio temporal sino también en el dominio espacial. Así, considerando M_R antenas en recepción suficientemente separadas, se puede conseguir a la salida de los M_R canales formados entre la antenna transmisora y las M_R antenas receptoras, M_R réplicas incorreladas de la misma señal transmitida. Una consecuencia de ello es que al combinar estas distintas réplicas, tal como indica la Figura 4.31, se puede conseguir a la salida de un combinador apropiado una nueva señal con un mayor nivel de potencia instantánea, y de este modo recuperar finalmente la señal de información modulada en mejores condiciones de relación potencia de señal a potencia de ruido.

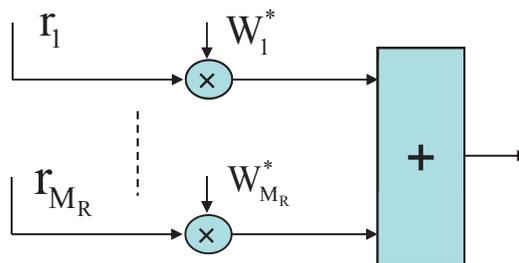


Figura 4.31 Esquema de Diversidad en Espacio

En estas circunstancias una combinación MRC (*Maximum Ratio Combining*) (Figura 4.31) es

una combinación lineal óptima donde los distintos pesos [8] $w_1^*, w_2^*, \dots, w_{M_R}^*$ resultantes conducen a que la relación señal a ruido resultante γ_R en cada instante de tiempo es la suma de las relaciones señal a ruido a la entrada del combinador $\gamma_1, \gamma_2, \dots, \gamma_{M_R}$:

$$\gamma_R = \gamma_1 + \gamma_2 + \dots + \gamma_{M_R} \tag{4.22}$$

La estructura de antenas anteriores se la conoce como SIMO (*Single Input Multiple Output*) y a la ganancia media obtenida en el valor de la relación señal a ruido γ se la denomina ganancia de la estructura de antenas o *array gain*. Al ser iguales las relaciones medias de señal a ruido en cada rama, resulta que el *array gain* es igual al número de antenas M_R .

La diversidad en espacio puede también hacerse en transmisión, de modo que la señal es transmitida por cada una de las M_T antenas después de haber sido convenientemente modificada en módulo y fase de modo que en recepción las distintas señales recibidas por la única antena receptora lleguen en fase y se sumen coherentemente [9]. Esta estrategia conduce a una estructura MISO (*Multiple Input Single Output*) y es una réplica MRC de la usada con la estructura SIMO, siendo ahora la relación señal a ruido en el receptor

$$\gamma = \gamma_1 + \gamma_2 + \dots + \gamma_{M_T} \tag{4.23}$$

y el *array gain* definido también como la ganancia media obtenida en el valor de la relación señal a ruido γ es ahora igual a M_T .

Con los esquemas en diversidad es factible a través de los combinadores adecuados [10] obtener comportamientos asintóticos de la tasa de error de bit en entornos Rayleigh de acuerdo con la expresión

$$P_b \rightarrow \frac{1}{\gamma_i^M} \tag{4.24}$$

donde γ_i es la relación señal a ruido en cada una de las ramas de recepción, supuestas iguales, y M el orden de la diversidad que coincide con el número de antenas. De una manera más formal se define la ganancia en diversidad como

$$-\frac{\log_2 P_b}{\log_2 \gamma} \tag{4.25}$$

que coincide con la ganancia en el valor de la pendiente de la tasa de error en función de la relación señal a ruido en una escala doble logarítmica obtenida para un sistema en diversidad con respecto a un sistema sin diversidad que tiene pendiente 1. Esta ganancia de diversidad, en conjunción con el *array gain*, caracteriza las prestaciones de estas estructuras de antenas.

Desde un punto de vista práctico la combinación MRC en transmisión no es simple de

conseguir, ya que es el receptor quien conoce las respuestas impulsionales de los M_T canales que se forman entre las M_T antenas transmisoras y la receptora, y este conocimiento se debe trasladar al transmisor vía señalización para su correcta implementación. En estas condiciones, las estructuras de combinación conocidas como técnicas ST (*Space –Time Diversity coding*) hacen esa señalización innecesaria a costa de una degradación en las prestaciones de la diversidad en lo que se refiere al *array gain*. Las técnicas usadas permiten extraer la diversidad inherente de los M_T canales formados en ausencia del conocimiento del canal en el transmisor. Las técnicas ST introducidas en [11] se basan en una codificación a través del espacio (distintas antenas) y el tiempo (símbolos sucesivos) [12] con objeto de conseguir la diversidad mencionada.

Tanto con la diversidad espacial en recepción, como con la diversidad espacial en transmisión, se pueden conseguir mejoras de la capacidad del canal resultante al aumentar la relación señal a ruido a la entrada del receptor, $SNR=\gamma$, linealmente con el número de antenas. No obstante ello, esta capacidad, siguiendo la conocida expresión introducida por Shannon, continúa teniendo la misma naturaleza de $C=B \log_2(1+\gamma)$ bits/s, inicialmente propuesta. Es decir la capacidad o máxima velocidad de transmisión C para una banda B que se podría conseguir libre de errores crece sólo de modo logarítmico con la relación señal a ruido.

Introduciendo en la formulación de la capacidad la relación comúnmente usada de energía de bit a densidad espectral de potencia de ruido (E_b/N_0), la presencia de M antenas ($M=M_T$ o $M=M_R$ dependiendo del tipo de diversidad) y en consecuencia un *array gain* = M , para conseguir una velocidad de transmisión R en bits por segundo (b/s), resulta

$$R < C = B \cdot \log_2 \left(1 + M \frac{E_b}{N_0} \frac{R}{B} \right), \quad (4.26)$$

o lo que es equivalente, la cota inferior de la E_b/N_0 necesaria para una velocidad R es

$$\frac{E_b}{N_0} \geq \frac{1}{M} \frac{2^{(R/B)} - 1}{R/B}. \quad (4.27)$$

La naturaleza exponencial de la anterior expresión nos indica que la relación $\frac{E_b}{N_0}$ puede hacerse muy elevada para valores de la velocidad de transmisión superiores a B , como es el caso en los sistemas que requieren una alta eficiencia espectral (medida como R/B b/s/Hz) como LTE. En tales sistemas, la diversidad espacial en recepción sola podría resultar insuficiente, aún contando con esquemas de transmisión avanzados que introduzcan codificación de canal y modulaciones adaptativas con distintos niveles, tal como QAM.

Una nueva estrategia derivada de explotar la presencia de M_T antenas en transmisión y M_R antenas en recepción surge con la aparición en los laboratorios Bell en USA [13] del concepto

de MIMO (*Multiple Input Multiple Output*). La diferencia, conceptualmente relevante en relación a las estrategias de diversidad SIMO y MISO, es que la estrategia MIMO pretende explotar la variación espacial del canal móvil creando de hecho L caminos de transmisión paralelos desacoplados (e.g., se envía información diferente por cada camino) entre el emisor y el receptor, siendo idealmente $L = \min(M_T, M_R)$. A este concepto se le conoce con el nombre de multiplexado espacial. Tal como detallaremos más adelante, la capacidad del canal en ciertas condiciones puede llegar a valer $C = L B \log_2(1 + \gamma)$. Observemos que en tales circunstancias la capacidad aumenta linealmente con el número de antenas, o lo que es equivalente, la E_b/N_0 necesaria para conseguir una velocidad de transmisión $R = C$ es:

$$\frac{E_b}{N_0} \geq \frac{2^{\frac{(R/B)}{L}} - 1}{R/B} \tag{4.28}$$

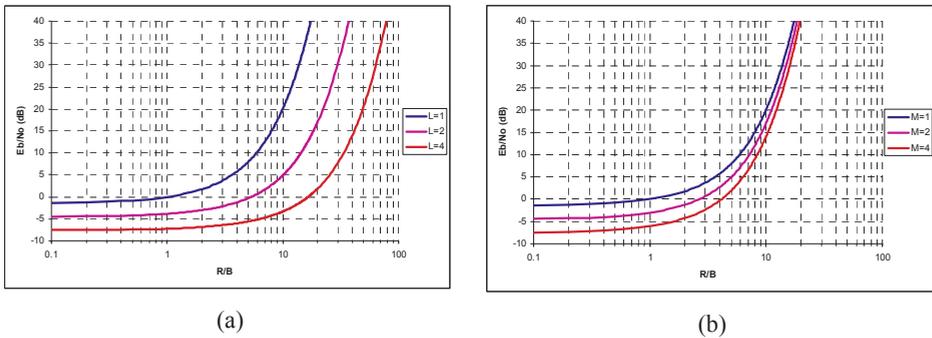


Figura 4.32 E_b/N_0 versus R/B para diferentes valores de (a) L en el caso de MIMO y (b) M en el caso de diversidad

A diferencia de los sistemas basados sólo en diversidad, la E_b/N_0 recibida necesaria para conseguir sistemas de alta eficiencia espectral, tales como los previstos en LTE, queda sustancialmente reducida al usar multiplexado espacial. En la Figura 4.32 se representa una cota inferior de la E_b/N_0 necesaria para distintos valores de eficiencia espectral medida como R/B . Se puede observar en la Figura 4.32 que a medida que aumenta L se pueden conseguir eficiencias espectrales más elevadas para los esquemas MIMO que para los esquemas que operan sólo con diversidad. En particular, obsérvese que para los sistemas en diversidad con eficiencias espectrales elevadas, la E_b/N_0 necesaria es mucho mayor que para los MIMO y es poco sensible al orden de la diversidad. Observemos también que para eficiencias R/B pequeñas o para E_b/N_0 pequeñas ambas estructuras, las MIMO y las de diversidad presentan comportamientos similares. En conclusión, las estructuras de antenas MIMO representan un salto cualitativo respecto a los sistemas de diversidad SIMO y MISO, y abren las puertas a conseguir sistemas de velocidad de transmisión elevadas, pero con potencias recibidas y por tanto coberturas y distancia entre emplazamientos abordables. Veremos a continuación una formulación para las distintas estructuras MIMO de interés en LTE, así como las capacidades de canal en cada caso. De hecho, debido a los extraordinarios avances logrados con la

conjunción de una codificación de elevadas prestaciones y modulaciones adaptativas ACM (*Adaptive and Coded Modulation*), la capacidad del canal en muchos casos no resulta un valor alejado del estado del arte, en cuanto a velocidades de transmisión alcanzables.

Una evolución de la estructura MIMO descrita anteriormente conduce a una distribución espacial de las antenas en un extremo de la estructura MIMO. Es decir, es posible vislumbrar que cada terminal móvil, que dispone de sólo una antena, es parte de una estructura distribuida de un MIMO que se comparte entre varios móviles. De esta manera varios usuarios no sólo comparten la estructura MIMO sino la banda de frecuencias con la que trabaja este MIMO distribuido, con las ventajas desde el punto de vista de eficiencia espectral que ello conlleva. Por otra parte, el MIMO deja de ser sólo un elemento de la capa de enlace y se incluye también como elemento a gestionar por el MAC del sistema, al tener que asignar en cada intervalo temporal qué terminales móviles forman parte de este MIMO distribuido y cuales no, dependiendo de las características del canal y de las prestaciones a nivel de retardo o *throughput* requeridas por cada usuario. A estas nuevas estructuras distribuidas se las llama *Multiple User MIMO*, en contraposición a las convencionales conocidas como *Single User MIMO*.

4.4.1 Caracterización de las estructuras MIMO

La introducción de las estructuras de antena MIMO levantó grandes expectativas que se han visto confirmadas en los últimos años por la multitud de publicaciones en este campo [14], y su implementación en sistemas concretos de comunicaciones móviles, como es el LTE, objeto de este libro. Es por ello que a continuación capturaremos los aspectos más relevantes de los MIMO necesarios para su comprensión, intentándolo conjuntar con una exposición detallada de sus particularidades técnicas más relevantes.

4.4.2 Formulación de las Estructuras MIMO

Podemos formular un sistema MIMO como un sistema de ecuaciones tal como se indica a continuación

$$y_i(t) = \sum_{j=1}^{M_T} h_{i,j}(\tau, t) * s_j(t), \quad i = 1, 2, \dots, M_R, \quad (4.29)$$

donde $y_i(t)$ es la señal recibida en el extremo receptor a la salida de la antena i -ésima, $h_{i,j}(\tau, t)$ es la respuesta impulsional del canal móvil en el instante t entre la antena transmisora j -ésima y la antena receptora i -ésima y $s_j(t)$ es la señal de entrada a la antena transmisora j -ésima. Podemos observar como esta formulación es una generalización de las estructuras SIMO (una antena en emisión y M_R en recepción) y MISO (M_T antenas en transmisión y una antena en recepción) antes mencionadas, en los que $s_j(t) = s(t)$ para todos los caminos. Esta formulación también captura la situación más convencional de las estructuras SISO (*Single Input Single Output*, una antena en emisión y una antena en recepción), en este

caso, el sistema de ecuaciones anterior se reduce a la formulación clásica para un canal móvil donde la señal de salida es la convolución de la señal de entrada y la respuesta impulsional del canal

$$y(t) = h(\tau, t) * s(t). \tag{4.30}$$

Nótese como la respuesta impulsional $h_{i,j}(\tau, t)$ es la de un sistema lineal variante en el tiempo, y dada la naturaleza indeterminada y variable del escenario móvil, $h_{i,j}(\tau, t)$ es también una señal aleatoria.

Como es usual en sistemas de ecuaciones lineales, una representación matricial de esta formulación conduce, no sólo a una representación más compacta de la misma, sino que permite una explotación de la teoría de matrices en el área de las estructuras MIMO.

Definimos entonces la matriz de dimensión $M_R \times M_T$ como

$$\mathbf{H}(\tau, t) = \begin{bmatrix} h_{1,1}(\tau, t) & h_{1,2}(\tau, t) & \dots & h_{1,M_T}(\tau, t) \\ h_{2,1}(\tau, t) & h_{2,2}(\tau, t) & \dots & h_{2,M_T}(\tau, t) \\ \vdots & \vdots & & \vdots \\ h_{M_R,1}(\tau, t) & h_{M_R,2}(\tau, t) & \dots & h_{M_R,M_T}(\tau, t) \end{bmatrix}. \tag{4.31}$$

En tal caso resulta,

$$\mathbf{y}(t) = \mathbf{H}(\tau, t) * \mathbf{s}(t), \tag{4.32}$$

donde

$\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_{M_T}(t)]^T$ es un vector de dimensión $M_T \times 1$ que contiene las señales de entrada $s_i(t)$ ($i=1, \dots, M_T$) a cada una de las M_T antenas transmisoras.

$\mathbf{y}(t) = [y_1(t), y_2(t), \dots, y_{M_R}(t)]^T$ es un vector de dimensión $M_R \times 1$ y contiene las M_R salidas de las M_R antenas receptoras.

* es el operador convolución y el superíndice T indica transpuesta.

En aras a simplificar la nomenclatura, a continuación definimos $\mathbf{y}(t)=\mathbf{y}$, $\mathbf{H}(\tau, t)=\mathbf{H}$ y $\mathbf{s}(t)=\mathbf{s}$. En tal caso resulta

$$\mathbf{y} = \mathbf{H} * \mathbf{s}. \tag{4.33}$$

En presencia de ruido térmico y para una potencia de señal transmitida en cada antena de valor P/M_T resulta.

$$\mathbf{y} = \sqrt{\frac{P}{M_T}} \mathbf{H} * \mathbf{s} + \mathbf{n}. \tag{4.34}$$

Consideramos el valor cuadrático medio de cada componente del vector \mathbf{s} igual a la unidad, de esta manera la potencia transmitida será P . A su vez, $\mathbf{n} = [n_1(t), n_2(t), \dots, n_{M_R}(t)]^T$ es un vector de ruido de dimensión $M_R \times 1$, siendo cada componente la contribución de ruido de la antena receptora correspondiente, estando todas las componentes incorreladas. A su vez, $n_i = n_{i,x} + jn_{i,y}$ ($i=1, \dots, M_R$), siendo $n_{i,x}$ y $n_{i,y}$ procesos estocásticos gaussianos incorrelados con densidad espectral de potencia N_0 .

Para un enlace caracterizado por una respuesta plana del canal, como puede ser el caso de OFDMA en relación a la banda estrecha ocupada por cada una de sus subportadoras, las entradas de la matriz \mathbf{H} son $h_{i,j}(\tau, t) = h_{i,j}(t) \delta(t)$. Los $M_T M_R$ posibles canales existentes entre las antenas transmisoras y receptoras son por tanto canales no dispersivos, aunque variantes en el tiempo y afectados por desvanecimientos. En tal caso podemos formular

$$\mathbf{y} = \sqrt{\frac{P}{M_T}} \mathbf{H} \mathbf{s} + \mathbf{n}, \quad (4.35)$$

donde \mathbf{H} es una matriz cuyas entradas son $h_{i,j}$.

Con objeto de visualizar cómo una estructura MIMO puede proporcionar $L = \min(M_T, M_R)$ caminos paralelos de propagación y el correspondiente multiplexado espacial, que es la esencia de las estructuras MIMO, hacemos uso de la descomposición en valores singulares SVD (*Singular Value Decomposition*) de la matriz de propagación \mathbf{H} , [15] de modo que

$$\mathbf{H} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H, \quad (4.36)$$

donde el superíndice H denota en la matriz correspondiente que es transpuesta y conjugada. Las matrices resultantes \mathbf{U} y \mathbf{V} son matrices de dimensión $M_R \times r$ y $M_T \times r$ respectivamente y satisfacen

$$\mathbf{U}^H \mathbf{U} = \mathbf{V}^H \mathbf{V} = \mathbf{I}_r \quad (4.37)$$

donde \mathbf{I}_r es la matriz identidad de dimensión $r \times r$, $\mathbf{\Sigma}$ es una matriz diagonal que presenta los valores $\{\sigma_1, \sigma_2, \dots, \sigma_r\}$ en la diagonal principal y cero en el resto de posiciones, lo que denotaremos como $\mathbf{\Sigma} = \text{diagonal} \{\sigma_1, \sigma_2, \dots, \sigma_r\} = \text{diagonal} \{\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_r}\}$, siendo σ_i ($i=1, \dots, r$) los valores singulares de la matriz \mathbf{H} .

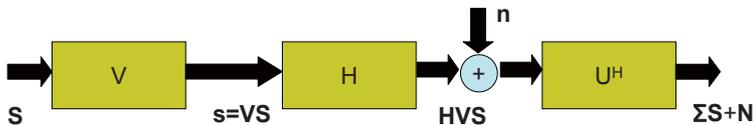


Figura 4.33 Descomposición SVD

El rango r de la matriz \mathbf{H} (definido como el número de ecuaciones linealmente independientes en el sistema de ecuaciones) nos establece el número de caminos paralelos de propagación que se pueden utilizar, cumpliéndose que $r \leq L = \min(M_T, M_R)$. Al valor de r también se le conoce como ganancia de multiplexado. Con objeto de ver la aplicabilidad de la descomposición SVD, introducimos un vector de entradas de señal de usuario $\mathbf{S} = (S_1, S_2, \dots, S_r)^T$ de dimensión $r \times 1$ que aplicamos a la matriz \mathbf{V} mediante la operación $\mathbf{V}\mathbf{S}$, que denominamos precodificación, tal como se esquematiza en la Figura 4.33. A la salida del denominado precodificador aparece el vector $\mathbf{s} = \mathbf{V}\mathbf{S} = (s_1, s_2, \dots, s_{M_T})^T$ de dimensión $M_T \times 1$ cuyos M_T valores constituyen las entradas a las M_T antenas transmisoras. De modo similar en recepción, y a resultados de la matriz de propagación \mathbf{H} , las M_R señales recibidas en las correspondientes M_R antenas receptoras coinciden con las componentes del vector $\mathbf{H}\mathbf{V}\mathbf{S}$ de dimensión $M_R \times 1$. Finalmente tal como se esquematiza en la Figura 4.33, aplicando este vector de salida a un decodificador caracterizado por \mathbf{U}^H resulta la ecuación

$$\mathbf{Y} = \sqrt{\frac{P}{M_T}} (\mathbf{U}^H \mathbf{H} \mathbf{V}) \mathbf{S} + \mathbf{U}^H \mathbf{n}. \tag{4.38}$$

Sustituyendo \mathbf{H} por su descomposición $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$, anteriormente definida, resulta

$$\mathbf{Y} = \sqrt{\frac{P}{M_T}} (\mathbf{U}^H \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H \mathbf{V}) \mathbf{S} + \mathbf{U}^H \mathbf{n} = \sqrt{\frac{P}{M_T}} \mathbf{\Sigma} \mathbf{S} + \mathbf{N}, \tag{4.39}$$

donde $\mathbf{N} = \mathbf{U}^H \mathbf{n}$ es un vector columna de r filas correspondiente al ruido resultante e \mathbf{Y} es el vector de salida. Es decir, a la salida de la operación de decodificación que determina el operador \mathbf{U}^H , tenemos r canales paralelos desacoplados

$$Y_i = \sqrt{\frac{P}{M_T}} \sigma_i S_i + N_i \quad (i = 1, \dots, r), \tag{4.40}$$

tal como se esquematiza en la Figura 4.34, donde se ilustra que a todos los efectos podemos imaginar que el esquema MIMO ha transportado las r señales a transmitir desde el extremo emisor al extremo receptor a través de r canales en paralelo desacoplados de ganancia $\sigma_1, \sigma_2, \dots, \sigma_r$ respectivamente.

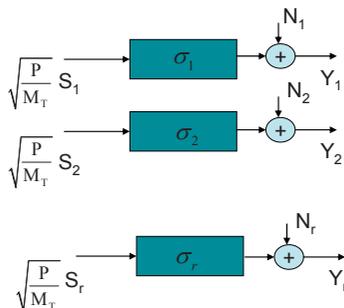


Figura 4.34 Esquema de una estructura MIMO

Cada columna de $\mathbf{V}^H \mathbf{V} = \mathbf{I}_r$ introduce un significado geométrico que nos permite una interpretación intuitiva del problema. En efecto, podemos considerar las columnas de la matriz \mathbf{V} como los r vectores unitarios de dimensión M_T ($\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_r$) que definen una base en un subespacio r -dimensional. La señal transmitida \mathbf{VS} se puede representar a partir de la base de este espacio vectorial según

$$\mathbf{VS} = S_1 \mathbf{e}_1 + S_2 \mathbf{e}_2 + \dots + S_r \mathbf{e}_r. \quad (4.41)$$

Los r vectores \mathbf{e}_i ($i=1, 2, \dots, r$) definen los r canales de transmisión ortogonales ($\mathbf{e}_i, \mathbf{e}_j = \delta_{ij}$) y por lo tanto desacoplados en los que cada emisor transporta las r señales de usuario recibidas, dando de esta forma sentido amplio al concepto de multiplexado espacial en contraposición al multiplexado temporal o frecuencial convencional usado en caso de tener canales ortogonales en el tiempo (TDM: *Time Division Multiplex*) o en la frecuencia (FDM: *Frequency Division Multiplex*), respectivamente.

Como se ha comentado, se cumple que $r \leq L = \min(M_T, M_R)$, de modo que L constituye el máximo número de caminos en paralelo que podemos extraer del canal móvil. Valores más pequeños que L son factibles, e incluso valores degradados de $r=1$ que indican que la estructura dispersiva de los canales móviles correspondientes no permite ninguna paralelización. Por otra parte, el vector \mathbf{S} debe ser tal que el valor medio de $\mathbf{S}^H \mathbf{S}$ es $E[\mathbf{S}^H \mathbf{S}] = M_T$ si pretendemos mantener que se cumpla que la potencia transmitida es P .

4.4.3 Capacidad de las Estructuras MIMO

La capacidad de un canal de comunicaciones que contemple estructuras con varias antenas transmisoras y receptoras se puede obtener generalizando la expresión clásica de la capacidad de un canal C de ancho de banda B y relación señal a ruido γ y introducida por primera vez por Shannon y ya mencionada anteriormente como

$$C = B \log_2(1 + \gamma). \quad (4.42)$$

El hecho de que las modernas técnicas de codificación y modulación permitan ya alcanzar velocidades de transmisión muy próximas al límite establecido por la capacidad de un canal de comunicaciones, otorgan 'si cabe' un mayor interés al conocimiento de esta expresión.

4.4.3.1 Canal conocido al transmisor

Cuando el canal es conocido al transmisor, esto es, la matriz \mathbf{H} es conocida al transmisor, podemos recurrir a la descomposición SVD de dicha matriz ilustrada anteriormente para obtener las matrices \mathbf{V} y \mathbf{U} de dicha descomposición. Es posible entonces precodificar las señales en el transmisor de acuerdo con el operador \mathbf{V} . La capacidad resultante es en tal caso la suma de la capacidad de los r canales desacoplados

$$y_i = \sqrt{\frac{P}{M_T}} \sqrt{\lambda_i} S_i + n_i, \quad i = 1, 2, \dots, r, \quad (4.43)$$

dando lugar a

$$C = B \sum_{i=1}^r \log_2 \left(1 + \frac{P}{M_T} \frac{\lambda_i}{BN_0} E \left[|S_i|^2 \right] \right), \quad (4.44)$$

donde $E [\]$ indica valor esperado, siendo

$$\frac{PE \left[|S_i|^2 \right]}{M_T} \quad (4.45)$$

el valor cuadrático medio de potencia asociado al subcanal i -ésimo y $\sigma_i = \sqrt{\lambda_i}$ la ganancia del subcanal i -ésimo. Dado que el canal \mathbf{H} es conocido al transmisor, también lo serán los valores de $\sigma_i = \sqrt{\lambda_i}$, siendo de este modo posible encontrar cual es el valor de $E \left[|S_i|^2 \right]$ óptimo, es decir la potencia de señal que maximiza la capacidad del canal, con la condición de que $E[\mathbf{S}^H \mathbf{S}] = M_T$. Este valor de potencia se puede encontrar a partir de un algoritmo iterativo *waterpouring algorithm* [16].

Obsérvese que la capacidad de un canal MIMO en estas condiciones no obedece a una fórmula cerrada y es en cualquier caso aleatoria en la medida que lo es el canal de propagación representado por \mathbf{H} . Es decir habrá unos entornos que favorecerán el multiplexado espacial más que otros. En todo caso corresponde a la mayor velocidad de transmisión que es posible transmitir con estructuras MIMO.

4.4.3.2 Canal desconocido al transmisor

Veamos ahora el caso de que no haya conocimiento del comportamiento del canal de propagación en el transmisor. En esta situación, el transmisor no puede conocer la matriz de precodificación \mathbf{V} , y en consecuencia, se consideran directamente las M_T señales de entrada $s_i(t)$ ($i=1, \dots, M_T$) a cada una de las M_T antenas transmisoras. Se asume que estas señales son independientes y presentan la misma potencia P/M_T , ya que, a priori, ante el desconocimiento del canal no hay otra estrategia de asignación de potencias mejor. Con objeto de determinar la capacidad en este caso supondremos a continuación una situación ilustrativa en un escenario particular con $M_T = M_R = M$ y que verifique que la matriz de propagación es ortogonal, esto es $\mathbf{H}\mathbf{H}^H = \mathbf{H}^H\mathbf{H} = M \mathbf{I}_M$ y por lo tanto su rango r es igual a M . Bajo estas circunstancias, la señal a la entrada del receptor vendrá dada por la expresión anteriormente mencionada como:

$$\mathbf{y} = \sqrt{\frac{P}{M_T}} \mathbf{H} \mathbf{s} + \mathbf{n}$$

donde las expresiones \mathbf{y} , \mathbf{H} , \mathbf{s} , \mathbf{n} coinciden con las descritas para la expresión (4.35), y con $E[|s_i|^2] = 1$, para que se satisfaga que la potencia total transmitida es P . En recepción, si se aplica la señal recibida y el operador \mathbf{H}^H , resulta

$$\mathbf{H}^H \mathbf{y} = \sqrt{\frac{P}{M}} \mathbf{H}^H \mathbf{H} \mathbf{s} + \mathbf{H}^H \mathbf{n}, \quad (4.46)$$

lo que conduce a las M ecuaciones

$$r_i = \sqrt{\frac{P}{M}} M s_i + N_i = \sqrt{PM} s_i + N_i, \quad (4.47)$$

donde $\mathbf{H}^H \mathbf{n} = (N_1, N_2, \dots, N_M)^T$ y

$$E[|N_i|^2] = E[N_i N_i^*] = ME[n_i n_i^*]. \quad (4.48)$$

La capacidad resultante, calculada como la agregación de la capacidad de los M canales desacoplados es

$$C = MB \log_2 \left(1 + \frac{PME [s_i s_i^*]}{BMN_0} \right) = MB \log_2 (1 + \gamma), \quad (4.49)$$

donde la relación señal a ruido es

$$\gamma = \frac{PE [s_i^2]}{BN_0} = \frac{P}{BN_0}. \quad (4.50)$$

Por otra parte, veamos ahora como en muchas situaciones prácticas la consideración de ortogonalidad que se ha tomado para la matriz \mathbf{H} puede constituir una buena aproximación. En efecto, supongamos un número elevado de antenas, con las distintas respuestas impulsionales h_{ij} siendo variables aleatorias gaussianas incorreladas para los distintos valores de i y j . Por la ley de los grandes números se cumpliría en estas circunstancias

$$\mathbf{H} \mathbf{H}^H = \mathbf{H}^H \mathbf{H} \rightarrow M \mathbf{I}_M. \quad (4.51)$$

Es decir, para M tendiendo a infinito podríamos poner el signo de igualdad en el límite de la expresión anterior, con la consiguiente validez de la capacidad calculada. De hecho, ya en muchas situaciones prácticas para M moderados, esta expresión puede ser una buena aproxi-

mación. Nótese por otra parte como la capacidad obtenida coincide con la utilizada anteriormente en el apartado 4.4.1, para obtener la velocidad de transmisión de un canal MIMO.

De un modo más formal, y sin ninguna suposición sobre la ortogonalidad de \mathbf{H} , se puede demostrar entonces que para una \mathbf{H} genérica [14] la capacidad del MIMO es

$$C = B \sum_{i=1}^r \log_2 \left(1 + \frac{\gamma}{M_T} \lambda_i \right) \tag{4.52}$$

donde $\sqrt{\lambda_i}$ coinciden con los valores singulares de la matriz \mathbf{H} . Observemos entonces como la capacidad de un sistema MIMO sin información de canal en el transmisor se comporta como la capacidad de r canales en paralelo cada uno con una ganancia de canal λ_i y potencia transmitida igual a la total dividida por M_T . En particular, el valor óptimo de la capacidad se cumple cuando la matriz de propagación es ortogonal [14], concluyéndose en tal caso que la capacidad de la estructura MIMO es M veces la capacidad de un enlace SISO como se ha obtenido antes en la expresión (4.49).

4.4.4 Precodificación

Tal como hemos visto anteriormente la capacidad máxima del sistema MIMO se obtiene cuando el emisor es capaz de conocer el valor \mathbf{H} de la matriz de propagación vista en el receptor, que se denomina “*Channel State Information in the Transmitter*” (CSIT).

Alternativamente, resulta suficiente si el emisor conoce la matriz \mathbf{V} resultante de la descomposición SVD de la matriz \mathbf{H} , para aplicar el proceso de precodificación. Desgraciadamente la obtención de esta información en el emisor resulta a un coste prohibitivo en términos de *overhead* o peaje de señalización. Nótese que los valores de la matriz \mathbf{V} que se obtendría en el receptor deberían transmitirse de modo que capturasen la variabilidad del canal móvil, es decir cada T segundos, siendo $T \ll \tau_c$, con τ_c el tiempo de coherencia del canal. A modo de ilustración con $T=0.1$ ms, $M_T=4$, $r=4$ y 8 bits de cuantificación por componente de la matriz \mathbf{V} , resultaría una velocidad de transmisión de $128 \cdot 10^4$ b/s. lo que sería impracticable.

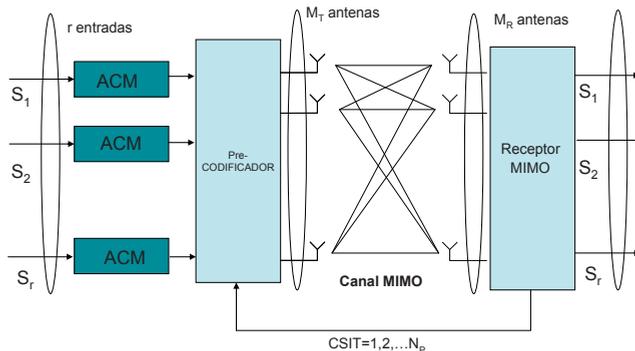


Figura 4.35 Estructura MIMO con Precodificación

En LTE con objeto de todavía disponer del estado del canal en transmisión (CSIT) se han habilitado estrategias subóptimas. Una de ellas es tener hasta un cierto número N_p de matrices de precodificación guardadas en transmisión, de manera que el receptor sólo tiene que enviar cual de las N_p matrices prerregistradas resulta la más conveniente. Esta estrategia reduce drásticamente el *overhead*, si bien a costa de aceptar un comportamiento con cierta degradación, pero aún aceptable [1]. Otros procedimientos similares son posibles y están ampliamente abordados en la literatura. Ciertamente con la paralelización en r canales, tal como se vio en la descomposición SVD de la matriz \mathbf{H} , cada canal está condicionado por una ganancia σ_i . Una optimización de la transmisión conduce entonces al uso de estrategias de modulación y codificación adaptativa que puede ser distinta en cada canal, de modo que los canales mejores que presenten σ_i elevadas tengan más niveles de modulación y menos redundancia de codificación y a la inversa para los canales peores que presenten σ_i pequeñas. La Figura 4.35 ilustra un esquema general donde se capturan las ideas anteriores en relación a la precodificación, la realimentación del CSIT y las técnicas ACM previas por las que los trenes de datos de usuario se codifican con códigos de canal apropiados (e.g turbocódigos) y se modulan con las estrategias nQAM adecuadas al estado del canal. En la medida que en un sistema real la precodificación no es perfecta, entonces la paralelización en r canales tampoco es ideal, y aparece una interferencia entre canales que debe eliminar, en la medida de lo posible, el receptor MIMO correspondiente. En la literatura se pueden encontrar descripciones con distintas estrategias de optimización, y usando bien receptores de tipo lineal o no lineal [1].

4.4.4.1 Codificación espacio-tiempo

El concepto de la codificación espacio-tiempo está asociado al uso de los distintos caminos de propagación que las diversas antenas de transmisión nos brindan para introducir la redundancia necesaria en un proceso de codificación, que en un sistema convencional implicaría un consumo de recursos del sistema en forma de incremento de ancho de banda. Con ello se pretende alcanzar unas mejores prestaciones en la tasa de error finalmente obtenida.

Una estructura pionera de codificación espacio-temporal, ampliamente usada en los sistemas UMTS, es la de Alamouti [12]. Este esquema no requiere conocimiento del canal y está indicada para explotar los distintos caminos de propagación habilitados en los sistemas MISO. Últimamente han aparecido otras muchas estructuras más elaboradas, aunque la inicial de Alamouti continúa siendo una referencia por su simplicidad.

Por otra parte, sabemos que no siempre es posible el poder hacer uso de las estructuras MIMO óptimas que presuponen un conocimiento del canal en el transmisor, como ocurre en situaciones de alta movilidad en las que el conocimiento del estado del canal en el transmisor no es viable. Además, puede haber ciertos servicios sensibles al retardo que no pueden permitirse el uso de *schedulers* para esperar a tener buenas condiciones de canal ya que pueden retardar excesivamente el acceso al canal del servicio. En estas circunstancias una buena opción es introducir el uso de la codificación espacio-tiempo en un sistema MIMO, que

podría incorporar además de la multiplicidad de caminos en transmisión, la diversidad espacial inherente en recepción. Veremos a continuación a modo ilustrativo el comportamiento de una estructura MIMO con $M_T=2$ y $M_R=2$, que utiliza una codificación espacio-tiempo en transmisión. Extensiones a esta estructura con un mayor número de antenas transmisoras y receptoras también son posibles.

La matriz de propagación \mathbf{H} para un canal con *fading* plano es

$$\mathbf{H} = \begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \end{bmatrix} \tag{4.53}$$

Hemos eliminado en la notación la variación temporal en aras de una mayor compactación en la formulación.

La señal transmitida a las dos antenas transmisoras en un intervalo temporal (*slot*) dado es s_1 y s_2 . De acuerdo con la codificación de Alamouti, en el *slot* sucesivo de señalización se entregan a las dos antenas transmisoras $-s_2^*$ y s_1^* tal como ilustra la Figura 4.36.

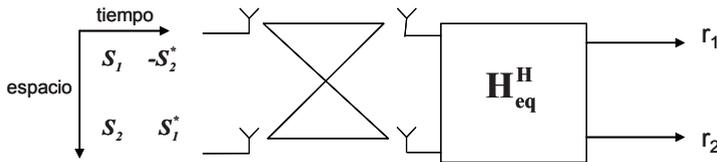


Figura 4.36 Esquema de Codificación de Alamouti

En estas circunstancias deducimos que

$$y_1 = \begin{bmatrix} y_{1,1} \\ y_{1,2} \end{bmatrix} = \sqrt{\frac{P}{2}} \begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}$$

$$y_2 = \begin{bmatrix} y_{2,1} \\ y_{2,2} \end{bmatrix} = \sqrt{\frac{P}{2}} \begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \end{bmatrix} \begin{bmatrix} -s_2^* \\ s_1^* \end{bmatrix} + \begin{bmatrix} n_3 \\ n_4 \end{bmatrix}$$

donde $y_{1,1}$ e $y_{1,2}$ corresponden a las señales presentes en las dos antenas receptoras durante el *slot* 1 y lo mismo para $y_{2,1}$ e $y_{2,2}$ durante el *slot* 2. De una manera más compacta podemos formular la anterior expresión como

$$\mathbf{y} = \begin{bmatrix} y_{1,1} \\ y_{1,2} \\ y_{2,1}^* \\ y_{2,2}^* \end{bmatrix} = \sqrt{\frac{P}{2}} \begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \\ h_{1,2}^* & -h_{1,1}^* \\ h_{2,2}^* & -h_{2,1}^* \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{bmatrix},$$

o bien

$$\mathbf{y} = \sqrt{\frac{P}{2}} \mathbf{H}_{\text{eq}} \mathbf{s} + \mathbf{n}, \quad (4.54)$$

donde

$$\mathbf{H}_{\text{eq}} = \begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \\ h_{1,2}^* & -h_{1,1}^* \\ h_{2,2}^* & -h_{2,1}^* \end{bmatrix},$$

$$\mathbf{s} = [\mathbf{s}_1, \mathbf{s}_2]^T,$$

$$\mathbf{y} = [y_{1,1}, y_{1,2}, y_{2,1}, y_{2,2}]^T,$$

$$\mathbf{n} = [n_1, n_2, n_3, n_4]^T.$$

Observando que la matriz \mathbf{H}_{eq} es ortogonal, con independencia de los valores de la matriz de propagación, es decir $\mathbf{H}_{\text{eq}}^H \mathbf{H}_{\text{eq}} = \|\mathbf{H}\|_{\text{F}}^2 \mathbf{I}_2$, donde \mathbf{I}_2 es una matriz identidad de dimensión 2 por 2 y $\|\mathbf{H}\|_{\text{F}}^2 = \sum_{i=1}^2 \sum_{j=1}^2 |h_{i,j}|^2$, la operación $\mathbf{r} = \mathbf{H}_{\text{eq}}^H \mathbf{y}$ y, sobre la señal recibida \mathbf{y} , nos conduce (ver Figura 4.36) a

$$\mathbf{r} = \sqrt{\frac{P}{2}} \|\mathbf{H}\|_{\text{F}}^2 \mathbf{I}_2 \mathbf{s} + \mathbf{n}', \quad (4.55)$$

donde \mathbf{n}' es el ruido gaussiano resultado de aplicar \mathbf{H}_{eq}^H sobre \mathbf{n} , es decir tenemos los dos valores

$$r_i = \sqrt{\frac{P}{2}} \|\mathbf{H}\|_{\text{F}}^2 s_i + n'_i, \quad (i=1,2)$$

para decidir los valores de los símbolos transmitidos s_1 y s_2 .

Observemos como ambos símbolos obedecen a la misma estructura de detección en las dos expresiones anteriores ($i=1,2$). Por otra parte, notemos también como la señal \mathbf{s} se ve multiplicada por un valor que es la suma de los módulos al cuadrado de los cuatro términos de ganancia del canal correspondiente, es decir tendremos un orden de diversidad equivalente de $M = M_T M_R = 4$, lo que incluye el producto de las diversidades del transmisor por las del receptor, que coinciden con el número de antenas correspondiente. En estas circunstancias la tasa de error de bit final correspondería asintóticamente a la de un sistema con ganancia de diversidad (orden de diversidad) de 4, tal como indica la expresión

$$P_b \rightarrow \frac{1}{\gamma^4}$$

La Figura 4.37 muestra resultados obtenidos con MATLAB para distintas configuraciones de estructuras de antenas considerando un entorno móvil con *fading* Rayleigh.

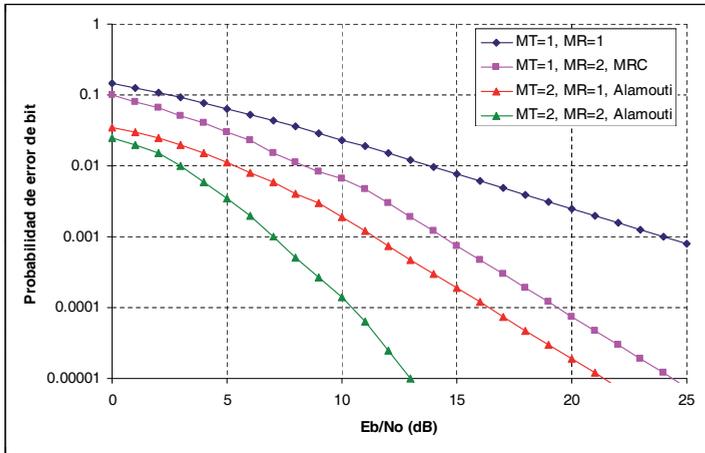


Figura 4.37 BER para distintas configuraciones de antenas

En el caso general con M_R antenas receptoras, la ganancia en diversidad resultante sería $2 M_R$. Extendiendo la codificación de Alamouti en transmisión obtendríamos en general ganancias de diversidad de $M_T M_R$ [14]. La aportación que introduce la estructura MIMO a la simple estructura MISO, en las que los códigos espacio-tiempo fueron primero introducidos, radica en el efecto multiplicativo de la ganancia de diversidad, lo que redundará en un decremento potencial de la tasa de error resultante.

Finalmente, suponiendo los valores medios $E \left[|h_{i,j}|^2 \right] = 1$, resulta un valor esperado: $E[\|\mathbf{H}\|_F^2] = 4$. En referencia al ruido la $E[n^n^{*H}] = 4N_0B$. En tal caso la relación potencia a señal recibida en el receptor es

$$\rho = \frac{\left[\|\mathbf{H}\|_F^2 \right]^2 P}{2 \|\mathbf{H}\|_F^2 N_0 B} = 2 \frac{P}{N_0 B} = 2 \gamma \quad (4.56)$$

Es decir, a diferencia de la ganancia en diversidad que es 4, el *array gain* es igual a 2, ya que solamente se puede extraer el *array gain* del receptor al no tener el transmisor conocimiento del canal.

Implementando esta estrategia observamos como podemos extraer la máxima diversidad dada por 4, sin embargo, al transmitir sólo un símbolo efectivo por *slot* (ya que de hecho se requieren dos *slots* para transmitir dos símbolos), la ganancia de multiplexado es $r=1$. Es decir el código de Alamouti maximiza la ganancia por diversidad pero no la ganancia de multiplexación. La pregunta que surge es cual es la mejor opción, el usar la estructura MIMO orientada a

diversidad con una ganancia de diversidad de $M_T M_R=4$ en el caso anterior, o por lo contrario, usar una estructura convencional orientada a conseguir una ganancia de multiplexado de 2. La respuesta no es inmediata y depende de las condiciones del canal. En [17] se ilustra una estrategia en diversidad para el caso de una modulación 16 QAM con el código de Alamouti, y otra estrategia de multiplexado espacial con dos modulaciones 4QAM y dos ramas en el supuesto de un CSIT perfecto en el transmisor. De esta manera, se comparan las dos estrategias con el mismo flujo de información entrante en la estructura de antenas y se establece un criterio para conmutar de una estrategia a la otra, observándose que no existe una estrategia mejor que otra en todas las circunstancias de propagación.

4.4.5 MIMO multiusuario (MU MIMO)

Una extensión en la explotación del multiplexado espacial surge con el MU MIMO (*Multi User MIMO*). A diferencia de las estructuras MIMO vistas anteriormente y que denominaremos SU MIMO (*Single User MIMO*), que competen solamente a la mejora de las prestaciones del enlace entre la estación de base y un móvil, habilitando para ello varios caminos paralelos desacoplados entre ambos, con el MU MIMO, lo que se pretende es que hasta K móviles distintos puedan compartir la misma banda de frecuencias. Se pretende con ello explotar una posible ortogonalidad o desacoplo espacial entre los flujos de datos que se transportan entre la base y cada uno de los distintos K móviles (Figura 4.38). En la medida que ello sea factible, lo que se consigue es ya no una mejora de velocidad de transmisión por unidad de banda en el enlace Base-Móvil, sino una mejor eficiencia en la suma de las velocidades de transmisión por unidad de banda en el conjunto de la celda. Es decir, con SU MIMO se pretende aumentar la capacidad del enlace mientras que con MU MIMO se pretende aumentar la capacidad de la celda. Cabe destacar que la primera *release* de LTE contempla estructuras MU MIMO, si bien con sólo una antena operativa en el receptor móvil.

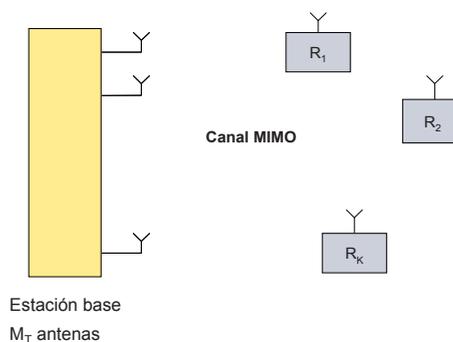


Figura 4.38 Esquema MIMO Multiusuario

4.4.5.1 MU MIMO en enlace descendente

En estas condiciones, podemos suponer todavía que una estructura MIMO con matriz de propagación \mathbf{H} existe, pero ahora coincidiendo M_R con el número de usuarios K que com-

parten la banda común asignada. Para el caso de que $M_T=M_R=K$, podemos formular la señal recibida por

$$\mathbf{y} = \sqrt{\frac{P}{M_T}} \mathbf{H}\mathbf{s} + \mathbf{n},$$

donde $\mathbf{s} = [s_1, s_2, \dots, s_{M_T}]^T$, $\mathbf{n} = [n_1, n_2, \dots, n_{M_T}]^T$, $\mathbf{y} = [y_1, y_2, \dots, y_{M_T}]^T$ y cada símbolo y_i corresponde al i -ésimo usuario. Con objeto de que un usuario no reciba como interferencia las señales dirigidas al resto de usuarios, se puede usar en transmisión una precodificación de los datos basada en la conocida estrategia de *zero forcing*. Es decir, en el supuesto que en emisión conozcamos la matriz de propagación \mathbf{H} , precodificamos previamente el vector de datos de usuario que denominamos $\mathbf{d} = [d_1, d_2, \dots, d_{M_T}]^T$ según,

$$\mathbf{s} = \mathbf{H}^{-1}\mathbf{d}, \tag{4.57}$$

con lo que en recepción obtenemos

$$\mathbf{y} = \sqrt{\frac{P}{M_T}} \mathbf{H}\mathbf{H}^{-1}\mathbf{d} + \mathbf{H}^{-1}\mathbf{n} = \sqrt{\frac{P}{M_T}} \mathbf{d} + \mathbf{H}^{-1}\mathbf{n}, \tag{4.58}$$

con lo que podemos recuperar el vector \mathbf{d} en recepción, es decir, se elimina la interferencia entre usuarios en el receptor, si bien a costa de aumentar el ruido para matrices de propagación mal condicionadas, que al invertirse pueden incrementar mucho el ruido final resultante. Esta estrategia por lo tanto puede ser satisfactoria sólo en situaciones de alta relación señal a ruido.

En el caso de que $M_R=K < M_T$ y, por lo tanto la matriz \mathbf{H} no sea simétrica, se puede llegar a una situación parecida pero ahora usando la pseudoinversa \mathbf{H}^\dagger como matriz de precodificación, dada por

$$\mathbf{H}^\dagger = \mathbf{H}^* (\mathbf{H} \mathbf{H}^*)^{-1} \mathbf{d}. \tag{4.59}$$

A la salida de de cada uno de los receptores tendríamos

$$\mathbf{y} = \sqrt{\frac{P}{M_T}} \mathbf{H} \mathbf{H}^* (\mathbf{H} \mathbf{H}^*)^{-1} \mathbf{d} + \mathbf{H}^* (\mathbf{H} \mathbf{H}^*)^{-1} \mathbf{n} = \sqrt{\frac{P}{M_T}} \mathbf{d} + \mathbf{H}^* (\mathbf{H} \mathbf{H}^*)^{-1} \mathbf{n},$$

con lo que podemos recuperar el vector \mathbf{d} en recepción.

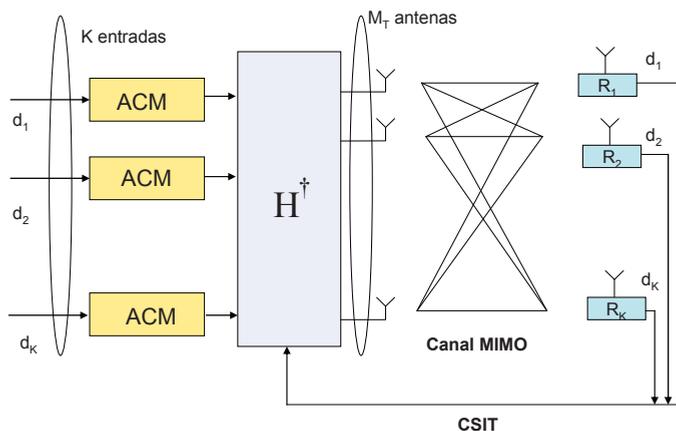


Figura 4.39 Estructura MU MIMO en enlace descendente

La Figura 4.39 captura una estructura MU MIMO para el enlace descendente con sólo una antena por receptor habilitada para la estructura MU MIMO, tal como se contempla en la *release 8* actual del LTE. De modo genérico, en esta figura se incorpora también para cada enlace con un usuario un esquema ACM de codificación y modulación adaptativa. En este caso, la matriz \mathbf{H}^\dagger (o \mathbf{H}^*) aplicaría al vector de salida de ACM que se decodificaría en los receptores correspondientes. Ciertamente, al igual como sucedía con la estructura MIMO para sólo un usuario, el conocimiento impreciso del canal en el receptor introduciría interferencia entre usuarios. Ello, unido a la poca robustez de la estrategia *zero forcing* en presencia de ruido, hace más apropiado el uso de estrategias MMSE, en que se parte de que va a existir interferencia entre usuarios y simplemente se minimiza con un criterio de error cuadrático mínimo. Se puede demostrar que en estas circunstancias un precodificador adecuado es [18]:

$$\mathbf{H}^* (\mathbf{H}\mathbf{H}^* + \alpha\mathbf{I})^{-1} \mathbf{d}. \quad (4.60)$$

Con $\alpha = K/P$ el valor que maximiza la relación señal a interferencia más ruido (*SINR*) en cada receptor y P la potencia total transmitida en la base.

Una problemática asociada al MU MIMO reside en que hay que elegir los K usuarios que comparten el enlace descendente en la misma banda de frecuencias y en el mismo tiempo. Es decir la base ha de ser informada de la matriz de propagación de todos los usuarios de la celda con anterioridad a la formación del grupo de K usuarios, cuyo valor está acotado a $(M_T)^2$ [19]. Nótese que el MU MIMO va mas allá de los aspectos de capa física que son los únicos que atañen a los SU MIMO y debe considerar además de los aspectos de potencia y velocidad de transmisión hacia cada terminal móvil, los aspectos relativos a la capa de MAC, al tener que decidir qué usuarios tienen acceso y cuales no en una cierta ventana temporal, en la que se supone que el canal de propagación no cambia significativamente.

Aunque sólo se haya considerado una antena operativa por móvil en el MU MIMO para

el enlace descendente, en futuras versiones del LTE se va a permitir más de una antena en recepción, de modo que el receptor móvil pueda explotar además de la diversidad espacial entre la base y los distintos móviles, el propio multiplexado espacial propio de los enlaces con SU MIMO. En este caso, dadas M_T antenas en transmisión y K usuarios, cada uno con M_R antenas en recepción (supongamos $M_R < M_T$), podrá haber hasta $M_R K$ caminos espaciales desacoplados. Se tratará entonces de encontrar cual es la configuración óptima de asignación de caminos a usuarios para cada cierta ventana temporal, con objeto de maximizar la velocidad de transmisión agregada que la base entrega a los K móviles. Nótese finalmente que el gran problema para este tipo de MU MIMO en el enlace descendente lo constituye el CSIT, que inevitablemente se traduce en carga de *overhead*.

4.4.5.2 MU MIMO en enlace ascendente

La realización de un MU MIMO, también conocido como MIMO MAC, es mucho más fácil de implementar que el anterior mencionado MU MIMO para el enlace descendente, en el que se requería que el transmisor conociese la matriz de propagación vista por el receptor (CSIT). La base en el MU MIMO implementa un clásico receptor multiusuario para separar las contribuciones de los K usuarios seleccionados. Para ello basta con que conozca la matriz de propagación en recepción obtenida según procedimientos convencionales de estimación de canal tal como se ilustra a continuación.

Representamos la señal recibida en la estación base, que consta de $M_R=M$ antenas, como un vector $\mathbf{y}=[y_1, y_2, \dots, y_M]^T$ dado por

$$\mathbf{y} = \sum_{i=1}^K \mathbf{h}_i s_i + \mathbf{n} = \mathbf{H}\mathbf{s} + \mathbf{n} \tag{4.61}$$

donde

- $\mathbf{h}_i = [h_{i,1}, h_{i,2}, \dots, h_{i,M}]^T$ es el vector cuyas componentes son las M respuestas impulsionales de cada usuario i -ésimo a cada una de las M antenas de la base, $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K]^T$,
- $\mathbf{s} = [s_1, s_2, \dots, s_K]^T$ es el vector de datos transmitidos por los K usuarios .
- \mathbf{n} es el vector de ruido térmico cuyas componentes son gaussianas e incorreladas.

Por otra parte en recepción, supuesto que la base conoce la matriz de propagación \mathbf{H} , o lo que es equivalente las respuestas impulsionales de cada usuario a cada una de las antenas, se procede a la detección del vector de datos \mathbf{s} transmitido. Un receptor representativo entre los varios existentes es el denominado ML (Maximum Likelihood), que resuelve la siguiente optimización, para la que existen algoritmos rápidos de ejecución:

$$\mathbf{s}' = \arg \min_{\mathbf{s}} \|\mathbf{y} - \mathbf{H}\mathbf{s}\|^2, \tag{4.62}$$

donde \mathbf{s}' es el vector de datos detectados.

4.4.6 Comparativa entre estructuras MIMO y MU MIMO

Las estructuras MU MIMO ciertamente presentan inconvenientes con respecto a las convencionales MIMO tales como:

- Efecto cerca-lejos (*Far-end*), ya que puede haber significativas diferencias en las pérdidas de propagación entre cada usuario y la base, que no siempre pueden ser compensadas debido a la máxima potencia transmitida por un móvil así como a los propios errores del control de potencia, lo que entre otras consecuencias puede redundar en la falta de equidad en el trato a todos los móviles por la base.
- El problema asociado a la propia gestión de la estructura MIMO hay que combinarlo con un *scheduler* adecuado, en relación a los usuarios que comparten un acceso a través de una firma espacial asignada a cada uno para poder habilitar un multiplexado espacial.
- La cooperación entre usuarios con distintas antenas es mucho más difícil que la cooperación entre las varias antenas que forman parte de una misma estructura.
- CSIT es un requerimiento exigente en el enlace descendente.

No obstante, las MU MIMO proporcionan también indiscutibles ventajas tales como:

- Proporcionan diversidad multiusuario, de modo que a través del *scheduler* se pueden elegir en cada ventana temporal los usuarios que presenten los mejores canales de propagación.
- La decorrelación entre antenas está habitualmente asegurada por la decorrelación de las firmas espaciales de los usuarios espacialmente separados.
- Permite multiplexado de usuarios en una misma banda de frecuencias además de la propia paralelización de los MIMOs.
- La posibilidad de que el rango de la matriz \mathbf{H} sea bajo, que puede suceder en los MIMO, es prácticamente despreciable con MU MIMO debido a la decorrelación espacial entre usuarios.
- Mitiga la necesidad de que los móviles tengan múltiples antenas, de hecho bastaría con una antena en los móviles para que una estructura MU MIMO fuera factible.

4.5 Estructuras de transmisión y recepción OFDMA con múltiples antenas

Este apartado pretende presentar la visión conjunta de la transmisión basada en OFDMA y en sistemas con múltiples antenas que se han presentado por separado en los apartados anteriores, en tanto que la combinación de ambos conceptos es la utilizada en la capa física de LTE.

Considérese un sistema de transmisión multiantena con un total de M_T antenas transmi-

oras y M_R antenas receptoras. Tal y como se ha comentado en el apartado 4.4.2, mediante este sistema es posible paralelizar un total de r canales MIMO donde r es el rango de la matriz \mathbf{H} que recoge las respuestas impulsionales entre las diferentes combinaciones de antenas transmisora y receptora, tal y como se mostró en la ecuación . De acuerdo con estas consideraciones, en la Figura 4.40 se muestra el esquema genérico de transmisión multiantena basado en OFDMA que permite la transmisión de N_F flujos de información.

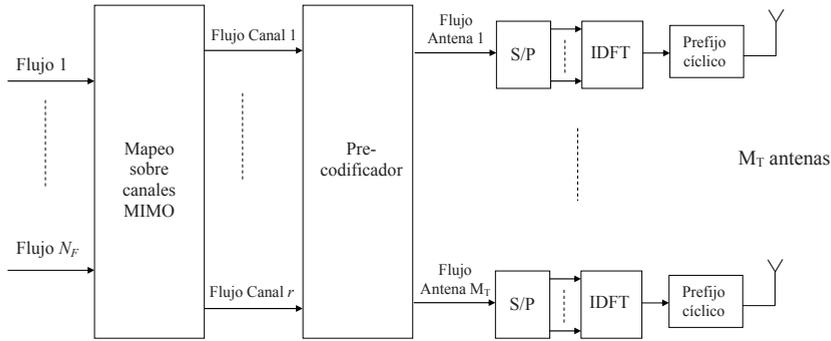


Figura 4.40 Esquema de transmisión OFDMA multiantena

Cada uno de los N_F flujos correspondería a un conjunto de símbolos modulados que constituyen la información a transmitir tras haber aplicado la codificación de canal. Dichos flujos pueden estar asociados por ejemplo a la información de diferentes usuarios o a diferentes flujos de información de un mismo usuario (e.g., diferentes canales de transporte).

Como puede apreciarse en la Figura 4.40, el procesado OFDMA, consistente en la IDFT y la inclusión del prefijo cíclico, tal y como se detalló en el apartado 4.1, se realiza por separado sobre cada una de las M_T antenas transmisoras, lo que se traduce en que sobre cada subportadora OFDMA se estarán enviando M_T símbolos simultáneamente, uno por cada antena. Así, previo al procesado OFDMA es preciso determinar cuales son los símbolos específicos que deben viajar sobre cada una de las antenas, lo que se lleva a cabo procesando los N_F flujos de información en dos fases que se detallan a continuación:

- **Mapeo sobre canales MIMO:** Este paso se encarga de determinar cómo se mapean los símbolos a transmitir de los diferentes flujos sobre los r canales MIMO existentes en paralelo, teniendo en cuenta que siempre debe existir un número de canales mayor o igual al número de flujos ($r \geq N_F$). En algunos casos este procedimiento es trivial, como por ejemplo cuando el número de flujos N_F coincide con el número de canales r , ya que en esta situación simplemente cada flujo se mapearía sobre un canal. Por el contrario, en el caso de disponer de un número de canales superior al número de flujos, $r > N_F$ pueden existir diferentes posibilidades de mapeo, dependiendo de cuantos canales se empleen para transmitir cada flujo, lo que permitiría conseguir velocidades de transmisión diferentes para cada flujo.

- Precodificación: En este paso se procesan los símbolos de cada uno de los r canales paralelos para obtener los símbolos que finalmente se deberán enviar a través de las M_T antenas transmisoras, de acuerdo con el mecanismo explicado anteriormente en los apartados 4.4.2 y 4.4.4. De forma genérica, este procesado se lleva a cabo multiplicando cada vector \mathbf{S} de r símbolos de entrada de dimensión $r \times 1$ por una matriz \mathbf{V} de dimensiones $M_T \times r$, lo que proporcionará a la salida de la operación $\mathbf{V}\mathbf{S}$ un vector de M_T símbolos cada uno de los cuales se enviará a través de una antena. Idealmente esta matriz \mathbf{V} debería obtenerse a partir de la estimación de la matriz \mathbf{H} que incluye el canal entre cada combinación de antena transmisora y receptora. Sin embargo, como ya se comentó en el apartado 4.4.4, con objeto de reducir la señalización que esto requeriría, es habitual que la matriz \mathbf{V} simplemente se escoja de entre un conjunto de matrices preestablecidas.

Es importante remarcar que la estructura de transmisión mostrada en la Figura 4.40 es general y aplicable a los diferentes tipos de transmisión multiantena, tanto si persiguen conseguir diversidad espacial en transmisión como si persiguen la multiplexación espacial de varios canales en paralelo. Por ejemplo, en el caso de utilizar diversidad espacial en transmisión mediante dos antenas, el esquema resultante sería el mostrado en la Figura 4.41. Como puede observarse, en este caso únicamente existe un flujo de símbolos a transmitir sobre $r=2$ canales en paralelo, de modo que el mapeo sobre los dos canales MIMO sería simplemente un conversor serie a paralelo (S/P). Análogamente, el proceso de precodificación sería en este caso el esquema de codificación de Alamouti que se presentó en el apartado 4.4.4 (ver Figura 4.36). Dicho esquema, por cada bloque de 2 símbolos que llegan en las dos entradas del precodificador (e.g., x_1, x_2 en la Figura 4.41), genera un conjunto de 4 símbolos, de los cuales $\{x_1, x_2\}$ se envían sobre una antena mientras que $\{-x_2^*, x_1^*\}$ se envían sobre la otra. En la figura se ilustra el proceso considerando los símbolos enviados sobre 4 subportadoras OFDMA. Como puede apreciarse, desde un punto de vista frecuencial, una subportadora transporta el símbolo x_1 por la primera antena y el símbolo $-x_2^*$ por la segunda, mientras que otra subportadora transporta el símbolo x_2 por la primera antena y el símbolo x_1^* por la segunda, por lo que realmente el proceso resulta en una codificación *espacio-frecuencial* (a diferencia del esquema de codificación original de Alamouti presentado en el apartado 4.4.4, que resultaba en una codificación *espacio-temporal* ya que los símbolos de una misma antena se enviaban en instantes diferentes de tiempo).

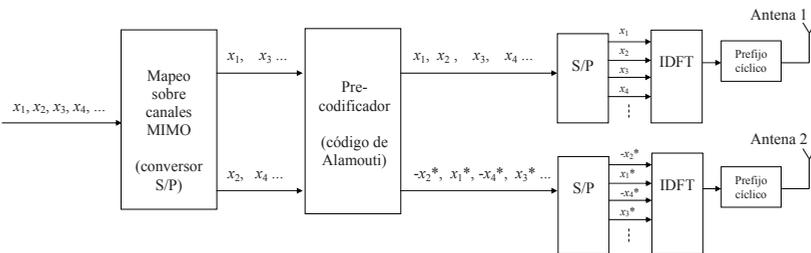


Figura 4.41 Transmisión MIMO/OFDMA para conseguir diversidad espacial con $M_T=2$ antenas

En la Figura 4.42 se muestra un ejemplo de transmisión MIMO/OFDMA para conseguir multiplexado espacial de $N_F=2$ flujos de información sobre $r=3$ canales en paralelo y empleando $M_T=4$ antenas transmisoras. En este caso, el mapeo sobre los tres canales MIMO disponibles se lleva a cabo utilizando un canal para el primer flujo $\{x_1, x_2, \dots\}$ y los otros dos canales para el segundo flujo $\{y_1, y_2, \dots\}$, de modo que este segundo flujo conseguirá enviar el doble de símbolos por unidad de tiempo que el primero.

El proceso de precodificación se lleva a cabo con una matriz \mathbf{V} de dimensiones 4×3 , multiplicando dicha matriz por cada vector de 3 símbolos \mathbf{S} de entrada al precodificador y obteniendo un vector \mathbf{zS} con los 4 símbolos que se enviarán sobre cada una de las antenas. Cada uno de los vectores de salida se envía sobre una subportadora diferente. Por ejemplo, el vector $\mathbf{z}(1)=[z_1(1), z_2(1), z_3(1), z_4(1)]^T$, resultante de la multiplicación de la matriz \mathbf{V} por el vector $\mathbf{S}_1=[x_1, y_1, y_2]^T$, contiene los 4 símbolos que se envían sobre la subportadora 1 en cada una de las cuatro antenas. Igualmente el vector $\mathbf{z}(2)=[z_1(2), z_2(2), z_3(2), z_4(2)]^T$, resultante de la multiplicación de la matriz \mathbf{V} por el vector $\mathbf{S}_2=[x_2, y_3, y_4]^T$, contiene los 4 símbolos que se envían a través de la subportadora 2.

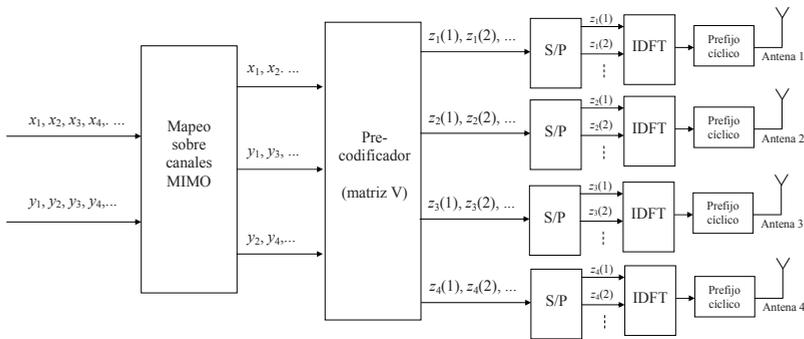


Figura 4.42 Ejemplo de transmisión MIMO/OFDMA con $N_F=2$ flujos, $r=3$ canales y $M_T=4$ antenas

En el caso concreto del enlace descendente de LTE, en la release 8 se soportan únicamente $N_F=2$ flujos [20], que pueden mapearse sobre $r=1, 2, 3$ ó 4 canales y que a su vez se traducen en $M_T=1, 2$ ó 4 antenas transmisoras. Nótese que no todas las combinaciones de N_F , r y M_T son válidas, ya que el valor de r está asociado con el número de antenas transmisoras M_T y receptoras M_R , debiendo ser r siempre menor o igual al mínimo de ambos valores, $r \leq \min(M_T, M_R)$, y además siempre deberá cumplirse que el número de canales sea mayor que el número de flujos a enviar, $r \geq N_F$.

En relación al proceso de recepción de la señal OFDMA multiantena, se muestra de forma genérica en la Figura 4.43 para una estructura con M_R antenas receptoras. Como puede apreciarse, los procesos son en esencia los inversos de la estructura del transmisor de la Figura 4.40. En concreto, después de efectuar el procesado OFDMA de la señal recibida por cada antena receptora (consistente en la extracción del prefijo cíclico y en realizar la DFT sobre los símbolos recibidos), se lleva a cabo el proceso de recepción MIMO para extraer la información

de cada uno de los r canales MIMO enviados en paralelo, a partir de los cuales se obtendrán los correspondientes N_F flujos enviados. Tal y como se detalló en los apartados 4.4.2 y 4.4.4, el proceso de recepción MIMO se lleva a cabo mediante la multiplicación del vector que contiene los r símbolos recibidos en cada antena por una matriz \mathbf{U}^H de dimensiones $r \times M_R$ que, conjuntamente con la matriz \mathbf{V} empleada en la precodificación del transmisor, constituyen idealmente la descomposición SVD de la matriz del canal \mathbf{H} , lo que permitiría la separación perfecta de las señales enviadas en cada canal. Sin embargo, como se comentó anteriormente, por cuestiones de reducción de señalización, las matrices utilizadas en la práctica no son las resultantes de la descomposición SVD sino que se emplean matrices predefinidas, por lo que existirá una cierta interferencia entre las señales de los diferentes canales, la cual se suele reducir incorporando en el receptor MIMO técnicas adicionales de procesamiento de señal.

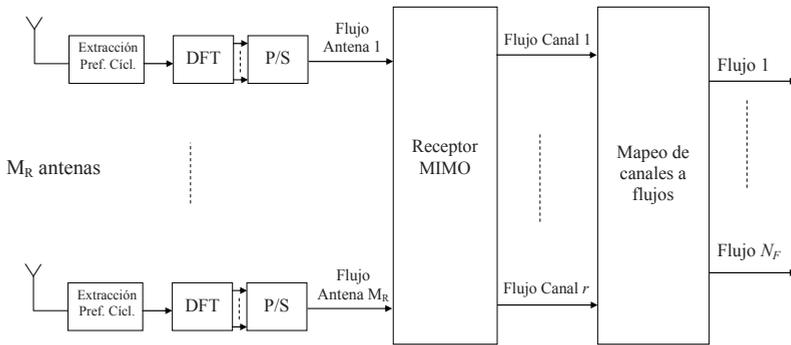


Figura 4.43 Esquema de recepción OFDMA multiantena

4.6 Referencias bibliográficas

- [1] F. Khan, *LTE for 4G Mobile Broadband*, Cambridge University Press, 2009.
- [2] E. Dahlman, S. Parkvall, J. Sköld, P. Beming, *3G Evolution. HSPA and LTE for Mobile Broadband*, Academic Press, Elsevier, 1a edición, 2007.
- [3] H. Sari, G. Karam, I. Jeanclaude, "Frequency-Domain Equalization of Mobile Radio and Terrestrial Broadcast Channels", IEEE Global Telecommunications Conference (GLOBE-COM), San Francisco, USA, 1994.
- [4] D. Falconer, S.L. Ariyavitakul, A. Benyamin-Seeyar, B. Eidson "Frequency Domain Equalization for Single-Carrier Broadband Wireless Systems", IEEE Communications Magazine, Vol. 40, No. 4, April, 2002, pp.58-66.
- [5] A. Pokhariyal, T.E. Kolding, P.E. Mogensen "Performance of Downlink Frequency Domain Packet Scheduling for the UTRAN Long Term Evolution", 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Helsinki, septiembre, 2006.

- [6] H.G.Myung, J. Lim, D.J. Goodman, "Single Carrier FDMA for Uplink Wireless Transmission", IEEE Vehicular Technology Magazine, septiembre, 2006, pp. 30-38.
- [7] O. Mauritz, B. M. Popovic, "Optimum Family of Spectrum-Shaping Functions for PAPR Reduction of DFT-Spread OFDM Signals", IEEE 64th Vehicular Technology Conference, VTC-2006 Fall, Montreal, Canada, Sept. 2006.
- [8] W.C.Y. Lee, *Mobile Communications Engineering*, Ed, McGraw Hill, 1982.
- [9] T.Lo, "Maximal Ratio Transmission", IEEE Transactions On Communications, October 1999.
- [10] W.C.Y. Lee, "Mobile Communications Design Fundamentals" Ed, Wiley Interscience, 1993.
- [11] Tarokh,V. Seshadri,N, Calderbank, A.R. "Space Time codes for High Data Rate Wireless Communications": Performance Criteria and Code Construction", IEEE Trans. On Information Theory, 1998.
- [12] S.Alamouti, "A Simple Transmit Diversity technique for wireless communication, IEEE Journal on Selected Areas in Communications, 1998.
- [13] G. Foschini, M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas", IEEE Wireless Personal Communications, March 1998.
- [14] A. Paulrag, R. Nabar, D. Gore "Introduction to Space-Time Wireless Communications" Cambridge University Press, 2003.
- [15] G. Golub , C.Van Loan, *Matrix Computation*, John Hopkins University press, 1989.
- [16] I. Talatar "Capacity of Multiple Antenna Gaussian Channels" European Trans. Tel., Nov/Dec 1999.
- [17] Robert W. Heath, Jr. and Arogyaswami J. Paulraj, "Switching Between Diversity and Multiplexing in MIMO Systems" IEEE Transactions on Communications, June 2005.
- [18] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multi-Antenna Multi-User Communication – Part I: Channel Inversion and Regularization" IEEE Transactions on Communications, Vol. 53, No.1, January, 2005, pp. 195-202.
- [19] D. Gesbert, M. Kountouris, R. W. Heath, Jr., C.-B. Chae, and T. Salzer, "Shifting de MIMO Paradigm: From Single User to Multiuser Communications", IEEE Signal Processing Magazine, vol. 24, no. 5, pp. 36-46, Oct., 2007.
- [20] 3GPP TS 36.211 V8.1.0 "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation", noviembre, 2007.

Anexo 4.1. Notación de señales complejas

Es habitual que en los sistemas de comunicaciones las técnicas de modulación empleadas exploten la ortogonalidad inherente entre señales sinusoidales desfasadas 90° para obtener dos canales de transmisión de la información simultáneos. De esta forma, se puede considerar que los símbolos de información enviados s constan de dos partes, la componente en fase, que se enviará modulada por un coseno, y la componente en cuadratura, que se enviará modulada por un seno. Esto admite una representación de los símbolos enviados según notación compleja como $s=I+jQ$ siendo I la componente en fase y Q la componente en cuadratura. La correspondiente señal modulada con una portadora a frecuencia f sería:

$$r(t) = I \cos(2\pi ft) - Q \sin(2\pi ft) \quad (4.63)$$

Esta última expresión puede formularse de modo más compacto como:

$$r(t) = \text{Re}\{s e^{j2\pi ft}\} \quad (4.64)$$

A partir de esta última expresión, y sobreentendiendo la parte real, es habitual expresar la señal modulada directamente en notación compleja como:

$$r(t) = s e^{j2\pi ft} \quad (4.65)$$

En la Figura 4.44 se muestra el diagrama de bloques correspondiente al proceso de modulación con señales reales, de acuerdo con (4.63), y su correspondiente equivalente con señales complejas, lo que da lugar a un modelo de representación más simplificado. Análogamente, en la Figura 4.45 se presentan los correspondientes modelos de demodulación del símbolo complejo s , realizados en este caso a partir de la multiplicación por la portadora correspondiente y por un filtro paso bajo (en la figura representado por un integrador a lo largo del período de símbolo T_s).

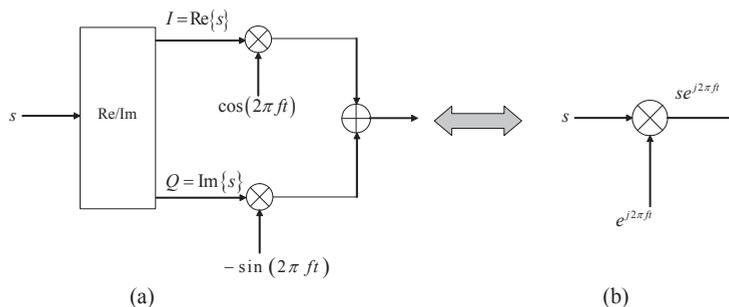


Figura 4.44 Modulación del símbolo complejo s . (a) Modelo con señales reales, (b) Modelo equivalente con señales complejas

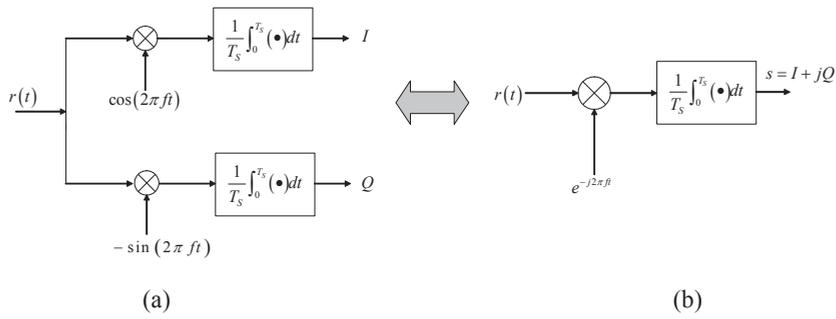


Figura 4.45 Demodulación del símbolo complejo s . (a) Modelo con señales reales, (b) Modelo equivalente con señales complejas



Interfaz Radio del Sistema LTE

5. Interfaz Radio del sistema LTE

Como se explica en el apartado 2.5.2.2, la interfaz aire permite la interconexión y el envío de tráfico y señalización entre el terminal móvil y las estaciones base denominadas *evolved NodeB* (eNB).

Los objetivos marcados por el organismo de estandarización 3GPP para el diseño de la interfaz aire del LTE son, [1]:

- Altas velocidades de transmisión de pico: mayores de 100 Mb/s en el enlace descendente y 50 Mb/s en el ascendente. Para ello se utilizarán modulaciones de elevada eficiencia espectral (por ejemplo 64 QAM en el enlace descendente).
- Mejora de la eficiencia espectral entre 2 y 4 veces comparada con la eficiencia espectral del sistema UMTS *versión (release) 6*. En particular se espera una eficiencia espectral mayor que 5bits/s/Hz utilizando modulaciones de alta eficiencia espectral y un método de acceso basado en la técnica OFDM.
- Mejora de la latencia (retardo) del sistema:
 - La latencia de la red de acceso radio en el plano de usuario debe ser menor de 10 ms.
 - Reducción significativa de la latencia del plano de señalización.

Ello se consigue dotando de inteligencia al eNB, es decir ubicando en él la mayor parte de las funcionalidades relacionadas con la gestión del acceso radio,

- Ancho de banda escalable: Valores estandarizados de 1.4, 3, 5, 10, 15 y 20 MHz.
- El sistema debe operar tanto en bandas pareadas (FDD) como en no pareadas (TDD).
- El sistema debe permitir/facilitar la interconexión con sistemas de comunicaciones móviles previos (UMTS, GPRS, etc.).
- El sistema LTE debe ser eficiente en costes,
 - Reducción de los costes de inversión (CAPEX) y de operación (OPEX),
 - Los costes también deben ser limitados en el caso de migración desde sistemas de comunicaciones móviles previos.

El objetivo de este capítulo es describir las principales funciones y características de la mencionada interfaz aire. Para ello, y a partir de la descripción de la pila de protocolos de dicha interfaz, se procederá a una descripción funcional de los diferentes protocolos asociados así como de los diferentes mecanismos de transferencia de información. Posteriormente se analizarán con detalle los mecanismos físicos utilizados para la transmisión/recepción tanto de la información de usuario como de la señalización, utilizando para ello el espectro radioeléctrico. Finalmente el capítulo termina describiendo algunos procedimientos básicos asociados a los mecanismos de capa física. El capítulo también contiene una serie de anexos

dedicados a la descripción de las propiedades de las secuencias utilizadas en la interfaz aire, descripción del denominado Sistema de Información de dicha interfaz, los mecanismos de re-transmisión híbridos en el sistema LTE y, del servicio *Multicast/Broadcast* (MBMS) y finalmente se procede a una descripción básica de los mecanismos de seguridad de dicha interfaz aire.

5.1 División funcional y Pila de protocolos

Como se ha indicado en el apartado 2.5.1, la red de acceso radio del sistema LTE está compuesta básicamente por eNBs (o estaciones base) a través de los cuales se conectan los terminales móviles a la red, utilizando para ello la denominada interfaz aire del sistema.

El eNB se conecta a la red troncal a través de la interfaz S1, que en realidad está desdoblada en dos interfaces diferentes: una para conectarse con el nodo encargado de las funciones de control en la red troncal (denominado *Mobility Management Entity o MME*) y otro que se encarga de las funciones de transferencia de los paquetes de usuario entre el eNB y en la red troncal a través del denominado *Server Gateway* (S-GW). Opcionalmente, los eNBs pueden conectarse entre si mediante la denominada interfaz X2 que permite, entre otras funciones, realizar la transferencia de paquetes de usuario entre eNBs durante el proceso de transferencia de llamada (*handover*) para mejorar sus prestaciones.

En la Figura 5-1 se resume la torre de protocolos de la interfaz aire del sistema LTE, inicialmente presentada en el apartado 2.5.3. Como es habitual en los equipos de telecomunicación, la estructura de esta torre de protocolos se articula en dos planos:

- El plano de usuario, destinado al envío de la información de usuario (paquetes IP generados por el terminal móvil o destinados a él).
- El plano de control o señalización, destinado a transportar los mensajes de control intercambiados entre la red y el terminal móvil, y que permiten un correcto funcionamiento del sistema.

Es importante destacar que el plano de usuario está interconectado al *Gateway* mediante los denominados Servicios Portadores S1 (i.e., *S1 Bearer*)¹ y por él se transmiten los mensajes de voz, datos o de radiodifusión y de grupo (*Broadcast/Multicast*). Por el contrario, el plano de control está interconectado al nodo MME, en donde están ubicadas, entre otras, las funcionalidades de gestión de la sesión y la movilidad y los mecanismos de control de las llamadas.

Desde el punto de vista de la interfaz aire, el envío de paquetes de usuario (paquetes IP) mediante dicha interfaz se gestiona en base al establecimiento de los denominados servicios portadores radio (i.e., *Radio Bearers*). Cada servicio portador radio tiene asociado un perfil de calidad de servicio (QoS) que la torre de protocolos que caracteriza a la interfaz debe garantizar.

¹ Véase sección 2.5.2.3 para la definición del concepto de portadora S1.

Los protocolos utilizados en el plano de usuario de la interfaz aire son los siguientes:

- *Packet Data Convergence Protocol* (PDCP) cuyas principales funciones son: compresión de las cabeceras de los paquetes IP, entrega/recepción ordenada de los paquetes IP desde/hacia las capas superiores, que están ubicadas en el *Gateway*.
- *Radio Link Control* (RLC), que permite establecer un enlace fiable a través de la interfaz aire.
- *Medium Access Control* (MAC) que permite un acceso ordenado de los diferentes usuarios (terminales móviles) al medio de transmisión (espectro radioeléctrico).
- Capa física, que proporciona el mecanismo físico de transmisión/recepción a través del espectro radioeléctrico.

Adicionalmente en la figura 5.1 también se muestra el denominado *Broadcast/Multicast Control Protocol* (BMC), que es el encargado de almacenar y gestionar la transmisión de los mensajes de grupo (*multicast*) o radiodifusión (*broadcast*) hacia el terminal móvil.

Por lo que respecta al plano de control, comparte con el plano de usuario los protocolos PDCP, RLC, MAC y capa física, e incorpora un nuevo protocolo denominado *Radio Resource Control* (RRC) que es quien se encarga de la gestión de la conexión a nivel radio entre el terminal móvil y la red (eNB), de la configuración de los servicios portadores, y del envío de mensajes relacionados con la movilidad desde/hacia el terminal móvil.

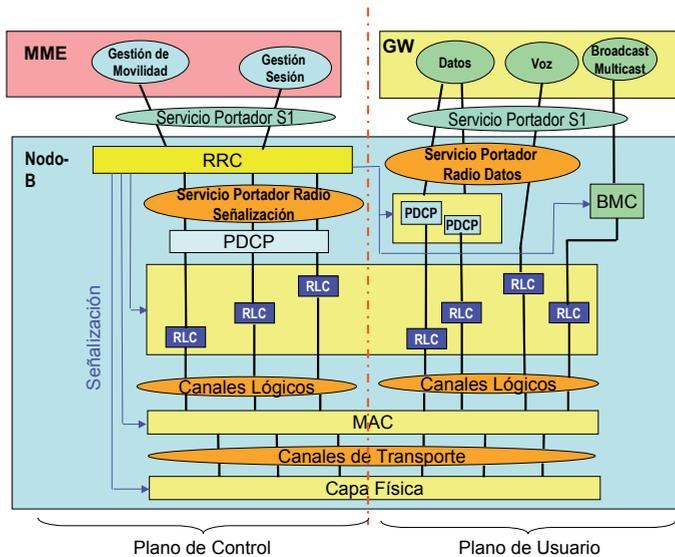


Figura 5-1 Torre de Protocolos de la interfaz aire del sistema LTE

En tanto que gestionan el comportamiento y las prestaciones del enlace radio, los protocolos PDCP, RLC y MAC se denominan protocolos de capa 2 (*Layer 2 protocols*) o protocolos de la capa de gestión del enlace. Por su parte el protocolo RRC se entiende como un protoco-

lo de capa 3 en tanto que gestiona funcionalidades a nivel de red. Obviamente el protocolo de la capa física es un protocolo de capa 1.

En las secciones posteriores de este capítulo se hace una descripción detallada de las funcionalidades básicas de los diferentes protocolos antes mencionados.

Por otra parte, la interrelación entre los diferentes protocolos se hace a partir de los denominados canales lógicos, de transporte y físicos. Los canales lógicos se establecen entre la capa RLC y la MAC y se utilizan para describir el tipo de información de usuario y/o señalización que se transmite a través de la interfaz aire. Los canales de transporte se establecen entre la capa MAC y la capa física y básicamente aportan flexibilidad a la interfaz aire, ya que permiten multiplexar diferentes canales lógicos en un canal de transporte. Además, los canales de transporte también se utilizan para describir cómo y con qué características se transmite la información a través de la interfaz aire. Finalmente los denominados canales físicos describen los mecanismos físicos de transmisión/recepción a través del enlace radio tanto de la información de usuario como de la propia señalización del sistema.

Por convención, desde el punto de vista de transmisión, a los paquetes que se entregan a un determinado nivel de la torre de protocolos se les denominan SDU (*Service Data Unit*) mientras que los paquetes que una capa entrega a la capa inferior del protocolo se les denominan PDU (*Protocol Data Unit*). Así por ejemplo, los paquetes que entrega, a través de los canales de transporte, la capa RLC hacia la capa MAC se les denominará RLC-PDU desde el punto de vista de la capa RLC y MAC-SDU desde el punto de vista de la capa MAC. En recepción el proceso es el inverso. Es decir, a los paquetes que se entregan a capas superiores se les denomina SDU, mientras que los que se reciben de capas inferiores se les denomina PDU.

Una característica importante de los SDUs y PDUs de la interfaz aire LTE es que están alineados a nivel de octeto o byte, es decir, el número de bits que componen los paquetes SDU o PDU siempre es múltiplo de ocho. Ello se hace así para facilitar la implementación de los protocolos y el manejo de los paquetes utilizando dispositivos micro-controladores o micro-procesadores.

5.2 Radio Resource Control (RRC)

El protocolo de Control de Recursos radio (*Radio Resource Control* o *RRC*), [2], es el encargado de controlar el comportamiento del móvil cuando éste está en el modo de operación "estado conectado", mientras que especifica la señalización de aviso (*paging*) y del Sistema de Información cuando el sistema todavía no tiene establecidos los servicios portadores radio (i.e., terminal en modo Idle).

Las principales funciones del protocolo RRC son:

- Radiodifusión de la información generada en el denominado Sistema de información ("*Information System*") (véase anexo 5.2), que incluye la difusión de información común (destinada a todos los terminales móviles) procedente de capas superiores (*Non*

Access Stratum o *NAS*), información aplicable cuando el terminal móvil está en modo *Idle* (por ejemplo: parámetros para la reelección de celda, información sobre celdas vecinas, etc.).

- Control de conexión RRC, que implica a todos los procedimientos relacionados con el establecimiento, mantenimiento y cierre de una conexión RRC, incluyendo los mecanismos de aviso (*paging*), establecimiento de las portadoras radio de señalización (*Signaling Radio Bearers* o *SRB*) y de las portadoras radio de datos de usuario (*Data Radio Bearers* o *DRB*), establecimiento inicial de los mecanismos de seguridad, integridad y cifrado, configuración de las capas inferiores (PDCP, RLC, MAC, física), etc.
- Gestión de movilidad entre tecnologías de acceso controladas por la red, incluyendo los procedimientos de movilidad, seguridad y transferencia de la información de contexto² por parte del terminal móvil.
- Gestión de los mecanismos de configuración de medidas y de transferencia de información en el caso de movilidad tanto entre tecnologías radio, como inter e intra frecuencia³.

Para la transferencia de los mensajes RRC así como de los mensajes de las capas superiores (*NAS*) se utilizan las denominadas portadoras radio de señalización (*SRB*). Se han especificado tres clases de *SRBs*:

- *SRB0* que se usa para transmitir mensajes RRC utilizando canales lógicos comunes de control.
- *SRB1* que se usa para transmitir mensajes *NAS* y la mayor parte de mensajes RRC. Se utilizan canales lógicos dedicados de control.
- *SRB2* que se usa para transmitir los mensajes RRC de alta prioridad. También se utilizan canales lógicos dedicados de control.

Todos los mensajes que utilizan canales lógicos dedicados de control⁴ utilizan mecanismos de protección de la integridad y cifrado de la información. También se utilizan mecanismos de retransmisión (*ARQ*) a nivel de capa RLC para garantizar un enlace radio fiable. Los mensajes que utilizan canales lógicos comunes de control no utilizan ninguno de los mecanismos mencionados anteriormente.

² La información de contexto es el conjunto de datos que contiene toda la información necesaria para establecer una sesión activa por parte de un móvil. Los datos incluyen, entre otros, la dirección IP asignada al usuario, los servicios a los que tiene acceso y la calidad de servicio asignada a los mismos.

³ Se denomina movilidad intra-frecuencia a los procedimientos de traspaso de llamada entre celdas que utilizan las mismas subportadoras, mientras que se denomina movilidad inter-frecuencia cuando las subportadoras de las celdas implicadas en el traspaso ocupan zonas del espectro diferentes.

⁴ Véase sección 5.4.1.1 para la definición de canales lógicos comunes y dedicados de control

5.3 Packet Dependence Convergence Protocol (PDCP): Funcionalidades Principales

Este protocolo ubicado sobre la sub-capa RLC aplica a todos los flujos de datos transmitidos en el plano de usuario así como a la mayoría de mensajes emitidos por el *Radio Resource Controller (RRC)* y transmitidos a través del plano de control de la interfaz aire LTE.

El protocolo PDCP, [3], tiene como funcionalidades más importantes:

- Compresión/descompresión de las cabeceras de los paquetes IP que llegan al eNB. El protocolo utilizado para ello es el denominado *Robust Header Compression (ROHC)* definido por el IETF en su estándar RFC3095, [4]. Mediante ésta compresión el tamaño de la cabecera se reduce de los 40 octetos típicos de una cabecera UDP/TCP-IP a un tamaño medio de cabecera entre 1 y 3 octetos.
- Cifrado de los datos y la señalización, utilizando para ello una clave de 128 bits.
- Implementa mecanismos de protección de la integridad de los mensajes de señalizados generados tanto por la capa RRC como por los protocolos de la parte de *Non-Access Stratum (NAS)* ubicados en la entidad de gestión de la movilidad (MME).
- Funcionalidades adicionales de capa 2 tales como reordenación y/o detección de duplicidades de paquetes RLC cuando se implementan mecanismos de movilidad entre eNBs.

Es importante subrayar que los mecanismos de cifrado y protección de la integridad son obligatorios en los flujos de señalización que se transmiten a través del plano de control de la interfaz aire LTE, mientras que el cifrado es opcional en los flujos de datos transmitidos a través del plano de usuario de dicha interfaz aire. Por lo que respecta al mecanismo de compresión de cabeceras, si bien teóricamente es obligatorio en todos los casos, hay que mencionar que el protocolo ROHC soporta un modo de operación *transparent uncompress mode* que permite mantener la cabecera del paquete IP inalterada.

5.4 Capa RLC-MAC

5.4.1 Funcionalidades relevantes de la sub-capa RLC

La subcapa RLC,[5], es la responsable de una transmisión fiable de la información a través de la interfaz aire. En particular:

- Implementa procedimientos de segmentación/concatenación de los paquetes IP recibidos de capas superiores (también denominados RLC_SDU) para adaptar su tamaño a las capacidades de transmisión de la interfaz aire. A los paquetes creados se les denomina RLC_PDUs.
- Implementa mecanismos de retransmisión de los RLC-PDU recibidos erróneamente, así como evita duplicidades y gestiona que los paquetes SDU_RLC recibidos se entreguen ordenados a las capas superiores.

Existen tres modos de operación de la sub-capa RLC:

- Modo Transparente (i.e., *Transparent Mode o TM*). Este es el modo de operación más simple ya que la sub-capa RLC entrega directamente a la sub-capa MAC los RLC_SDU recibidos de las capas superiores. Es decir, no se introduce ninguna cabecera a nivel de capa RLC, ni existen mecanismos de segmentación/concatenación, ni de retransmisión. Este modo de operación es típico de los canales lógicos de radiodifusión (BCCH) o de aviso (PCCH).
- Modo de No-reconocimiento (i.e., *Unacknowledged Mode o UM*). Este modo de operación detecta si los paquetes SDU_RLC recibidos contienen errores, pero no implementa mecanismos de retransmisión. También realiza funciones de segmentación y reensamblado así como gestiona la reordenación, si es necesario, de los SDU_RLC recibidos. Para ello utiliza un campo de la cabecera RLC denominado número de secuencia (i.e., *Sequence Number*). Este modo de operación fundamentalmente aplica a los canales de tráfico (TCH) y a canales *multicast* (MTCH), dependiendo del tipo de aplicación y de la calidad de servicio requerida (QoS).
- Modo de Reconocimiento (i.e., *Acknowledged Mode o AM*). Es el modo de operación más robusto. Además de las funciones de segmentación/reensamblado y de ordenación de los RLC-SDU también proporciona mecanismos de retransmisión tipo ARQ para eliminar posibles errores introducidos por el canal radio. Este modo de operación típicamente aplica a los canales dedicados tanto de tráfico (DTCH), dependiendo de la calidad de servicio requerida (QoS), como de control (DCCH).

Existe un mapeo directo entre los flujos de información y/o control recibidos por la sub-capa RLC y los canales lógicos definidos entre las capas RLC y MAC.

5.4.1.1 Canales Lógicos

Los canales lógicos se utilizan para describir el tipo de información que se transmite a través de la interfaz aire. Los canales lógicos se clasifican en canales lógicos de tráfico y canales lógicos de control.

Canales Lógicos de Tráfico

- *DTCH (Dedicated Traffic Channel)*. Es un canal punto a punto destinado a transferir información entre la red (eNB) y un terminal móvil determinado. Puede ser utilizado para transferir información de usuario así como señalización a nivel de aplicación asociada a un flujo de datos de usuario. Este canal lógico existe tanto en el enlace descendente como ascendente.
- *MTCH (Multicast Traffic Channel)*. Es un canal punto-multipunto utilizado para transmitir desde la red (eNB) información a un grupo de usuarios. Típicamente este canal

se utiliza para implementar el servicio MBMS (*Multimedia Broadcast and Multicast Service*)⁵. Este canal lógico sólo existe en el enlace descendente.

Canales Lógicos de Control

- *Broadcast Control Channel (BCCH)*. Este es un canal de control que utiliza el eNB para transmitir a los terminales móviles bajo su área de cobertura información sobre el sistema (e.g., la identificación del operador y la celda, configuración de los canales comunes de control o como acceder al sistema). Este es un canal lógico que sólo existe en el enlace descendente.
- *Paging Control Channel (PCCH)*. Este es un canal que utiliza la red (eNB) para llamar a terminales móviles de los que se conoce cual es su área de localización pero no la celda o eNB en la que están ubicados. Como en el caso anterior, este canal lógico sólo existe en el enlace descendente.
- *Common Control Channel (CCCH)*. Este canal permite la comunicación entre el eNB y el terminal móvil cuando todavía no se ha establecido una conexión a nivel de RRC. Se utiliza para enviar información de control fundamentalmente en las fases iniciales del establecimiento de la conexión. Este canal lógico existe tanto en el enlace descendente como ascendente.
- *Dedicated Control Channel (DCCH)*. Es un canal punto a punto destinado a transferir información de control entre la red (eNB) y un terminal móvil determinado, una vez que se dispone de una conexión a nivel de RRC. Es importante subrayar que este canal lógico sólo contiene información de control procedente del RRC y señalización a nivel de NAS, pero no señalización a nivel de aplicación asociada a un flujo de datos de usuario. Este canal lógico existe tanto en el enlace descendente como ascendente.
- *Multicast Control Channel (MCCH)* Es un canal punto-multipunto utilizado para transmitir información de control desde la red (eNB) a un grupo de usuarios que reciben servicios MBMS. Este es un canal lógico que sólo existe en el enlace descendente.

5.4.2 Funcionalidades relevantes de la sub-capa MAC

La sub-capa MAC, [6], proporciona el enlace entre los servicios proporcionados por la sub-capa RLC y la capa física. La sub-capa MAC proporciona las siguientes funcionalidades:

- Mecanismos de corrección de errores mediante procedimientos de retransmisión (Hybrid ARQ o HARQ). Este es un mecanismo de retransmisión, distinto del situado en la capa RLC, que se aplica ahora sobre los canales de transporte. El procedimiento HARQ utiliza técnicas de retransmisión basadas en el incremento de redundancia. Esto es, inicialmente se transmite la información contenida en canal de transporte con un mínimo de redundancia capaz de corregir los errores que introduce el canal radio, si éste no está fuertemente desvanecido. Si en recepción con esta redundancia no basta

⁵ Véase anexo 5.4 de este capítulo

para corregir los errores, en lugar de volver a retransmitir todo el paquete, como en el caso de las técnicas de retransmisión convencionales, ahora se solicita la transmisión sólo de información de redundancia adicional, no transmitida anteriormente. Una vez recibida esta información de redundancia adicional, se añade a la ya disponible en el receptor para volver a intentar detectar correctamente la información transmitida por el canal de transporte. Si nuevamente se produce una detección errónea de la información transmitida se repite el proceso hasta que o bien se recibe correctamente dicho canal de transporte o bien se completa la transmisión de toda la información de redundancia asignada al código utilizado. Con este mecanismo de transmisión parcial de la redundancia de codificación se busca optimizar el uso del canal radio, transmitiendo la mínima información de redundancia posible para que en recepción se pueda detectar correctamente el canal de transporte transmitido.

- Gestión de prioridades entre canales lógicos de un mismo terminal móvil así como entre terminales móviles, utilizando técnicas de gestión dinámica de recursos (*Dynamic Scheduling*).
- Selección del formato de transmisión (tipo de modulación, tamaño del bloque de transporte, etc.) a utilizar por parte de la capa física.
- Multiplexado/demultiplexado de paquetes MAC (MAC_SDUs) pertenecientes a diferentes canales lógicos en los denominados bloques de transporte (*Transport Blocks o TB*) para ser entregados/recibidos a/de los canales físicos utilizando para ello los denominados canales de transporte. En el caso que existan diferentes posibilidades de mapeo entre los canales de transporte y los canales físicos, la sub-capa MAC selecciona la configuración más apropiada, tomando en consideración las posibles configuraciones definidas por el operador.
- Informar sobre la gestión de los recursos radio. De forma general, el sistema LTE gestiona los recursos de dos formas: gestión dinámica y gestión semi-estática o persistente. En el caso de gestión dinámica de recursos, utiliza el enlace descendente (eNB hacia Móvil) para enviar mensajes de asignación de recursos y el enlace ascendente (Móvil hacia eNB) para confirmar dicha asignación, subtrama a subtrama⁶. En el caso de asignación semi-estática el procedimiento es similar, sólo que los recursos asignados en el enlace ascendente lo son por un periodo de tiempo mayor que una subtrama. En ambos casos el móvil envía al eNB los denominados *Buffer Status Reports* que informan del número de bytes almacenados en las colas (*buffer*) del móvil todavía pendientes de transmitir. A partir de esta información el eNB puede realizar la asignación de recursos más apropiada, tanto en el enlace ascendente como descendente, en función de los requisitos de calidad de servicio asociados a cada flujo de información (ó servicio establecido).

⁶ Véase sección 5.5.2 para la definición del concepto de subtrama. En todo caso comentar que la duración de una subtrama en el sistema LTE es 1 ms

5.4.2.1 Canales de Transporte

Los canales de transporte se utilizan para describir cómo y con qué características se transmite la información a través de la interfaz aire. Los canales de transporte se clasifican en función de que se utilicen en el enlace descendente o ascendente:

Enlace descendente :

- *Broadcast Channel (BCH)*. Este canal de transporte está unívocamente asociado con el canal lógico BCCH. Tiene un formato fijo, y se transmite en toda el área de cobertura del eNB.
- *Downlink Shared Channel (DL-SCH)*. Transporta tanto información de usuario como señalización. Soporta técnicas de retransmisión basadas en HARQ, y adaptación dinámica de enlace radio (i.e., *dynamic link adaptation*) variando modulación y códigos. Permite que le sean asignados recursos radio tanto de forma dinámica como semi-estática, así como la posibilidad de implementar mecanismos de recepción discontinua en el terminal móvil (DRX) para aumentar la eficiencia en potencia. En el caso del servicio MBMS, éste canal se puede configurar para difusión en toda la celda o bien en parte de ella utilizando técnicas de conformación de haz (*beamforming*).
- *Paging Channel (PCH)*. Este canal se transmite en toda el área de cobertura de la celda. Permite llamar a los terminales móviles así como informarlos de actualizaciones realizadas en el denominado "Sistema de Información". Permite una asignación dinámica de recursos.
- *Multicast Channel (MCH)*. Transmitido en toda el área de cobertura de la celda. En el caso del servicio MBMS, facilita el uso de una red iso-frecuencia (i.e., *Single Frequency Network*)⁷ utilizando una trama temporal con prefijo cíclico largo y una gestión semi-estática de los recursos radio.

Enlace ascendente

- *Uplink Shared Channel (UL-SCH)*: Transporta tanto información de usuario como señalización. Soporta técnicas de adaptación dinámica del enlace (modulación, codificación, potencia de transmisión.) así como mecanismos de retransmisión basados en HARQ. Permite la asignación de recursos radio tanto de forma dinámica como semi-estática. Está preparado para permitir el uso de técnicas de conformación de haz (i.e., *beamforming*).
- *Random Access Channel (RACH)*: Se utiliza para enviar peticiones de acceso al sistema, así como para transmitir mensajes cortos de información de control. Al ser un canal compartido por los diferentes usuarios del sistema existe riesgo de colisiones.

⁷ Véase sección 4.2.2 del capítulo anterior.

5.5 Capa Física

La capa física de la interfaz aire del sistema LTE, [7], se basa en la utilización de técnicas de acceso múltiple OFDMA en el enlace descendente (eNB a Móvil) y SC-FDMA en el enlace ascendente (Móvil a eNB). En ambos casos la separación entre subportadoras es fija e igual a 15 KHz. El número de subportadoras disponibles varía en función de la canalización del sistema, tal como se indica en la siguiente tabla.

Tabla 5-1 Canalización en el sistema LTE

Canalización	1,4 MHz	3 MHz	5 MHz	10 MHz	15 MHz	20 MHz.
Tamaño FFT	128	256	512	1024	1536	2048
Número de subportadoras disponibles	73	181	301	601	901	1201

La capa física del sistema LTE está pensada para operar en bandas altas de UHF, por encima de los 450MHz y hasta los 3,5 GHz, [8]. De hecho el estándar define hasta 40 posibles bandas de operación⁸ trabajando bien en modo de duplexión por división de frecuencia (FDD) o bien en duplexión por división en el tiempo TDD. La mínima separación entre posibles frecuencias portadoras, denominada *Channel Raster*, es de 100 KHz, de modo que la frecuencia portadora de un radiocanal viene dada por:

$$F_{DL} = F_{DL_low} + 0.1 \left[N_{DL} - N_{offs-DL} \right]$$

$$F_{UL} = F_{UL_low} + 0.1 \left[N_{UL} - N_{offs-UL} \right]$$

donde N_{DL} y N_{UL} representan los denominados *Absolute Radio Frequency Channel Number (ARFN)* del enlace ascendente y descendente mientras que $N_{offs-DL}$ y $N_{offs-UL}$ representan unos valores Offset (también asociados al enlace ascendente y descendente) que dependen de la banda de operación considerada. Por su parte F_{DL_low} y F_{UL_low} son las frecuencia de los límites inferiores de la banda de operación considerada. En la referencia [8] están tabulados los valores de $N_{offs-DL}$ y $N_{offs-UL}$ así como los posibles rangos de valores de N_{DL} y N_{UL} para cada una de las posibles bandas de operación del sistema. Es importante destacar que, dada una banda de operación, no existe un valor fijo de la frecuencia portadora sino un rango de posibles valores con una separación entre ellos igual al valor del *Channel Raster*.

Los posibles esquemas de modulación son: QPSK, 16QAM y 64QAM para el enlace descendente (DL) y QPSK y 16 QAM para el enlace ascendente (UL). Opcionalmente se contempla la posibilidad de incluir 64 QAM en el enlace ascendente en función de las capacidades del terminal móvil.

⁸ Véase tabla 6.2 en el capítulo sexto para una completa descripción de los límites superiores e inferiores de configuración de cada una de las bandas indicadas en el estándar.

Utilizando técnicas MIMO (2x2) y para una canalización de 20 MHz es posible alcanzar una velocidad de transmisión de pico a nivel de capa física de 150 Mb/s en el enlace descendente y 75 Mb/s en el ascendente.

5.5.1 Concepto de bloque de recursos físicos (Physical Resource Block)

Se denomina PRB (*Physical Resource Block*), al mínimo elemento de información que puede ser asignado por el eNB a un terminal móvil. Un PRB ocupa 180 KHz de banda equivalente a doce sub-portadoras equi-espaciadas 15KHz entre ellas, y en él se transmiten 6 ó 7 símbolos OFDMA, dependiendo de la longitud de prefijo cíclico utilizada. La duración de un PRB es igual a 0,5 ms, es decir la duración de un slot o ranura temporal.

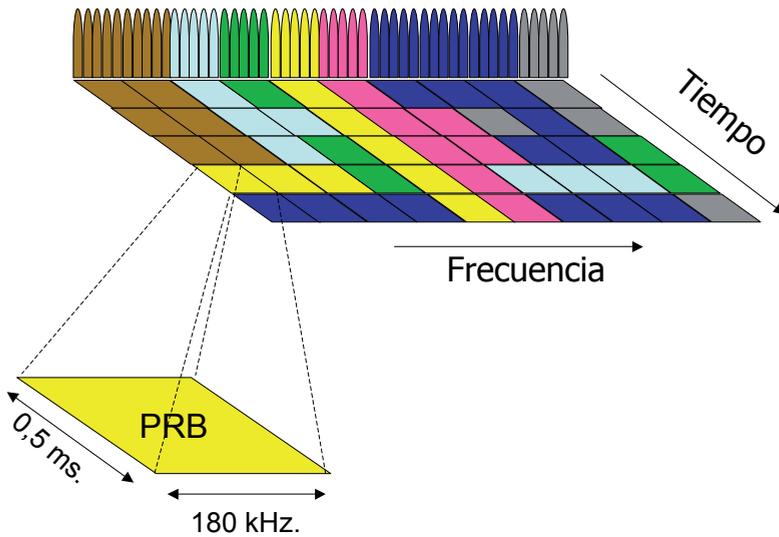


Figura 5-2 Concepto de Radio Bloque (*Physical Resource Block*)

El número de bloques de recursos físicos disponible en el sistema depende de la canalización tal como se indica en la siguiente tabla:

Tabla 5-2 Número de PRBs en función de la canalización

Canalización	1,4 MHz	3 MHz	5 MHz	10 MHz	15 MHz	20 MHz.
Número de PRB	6	15	25	50	75	100

Hay que subrayar que el número de subportadoras disponibles (indicado en Tabla 5-1) y el número de PRB (indicado en la Tabla 5-2) están relacionados. En efecto el número de subportadoras es 12 veces el número de PRBs más una, ya que hay que considerar que la subportadora central (ó subportadora DC) de la banda disponible no se utiliza para transmitir

información. La subportadora DC se utiliza para facilitar los mecanismos de ajuste y sincronización en frecuencia del receptor.

Ya que en un bloque de recursos físicos (PRB) disponemos de 7 símbolos con 12 subportadoras asociadas a cada uno de ellos, tenemos en total 84 recursos disponibles en donde ubicar los símbolos QPSK, 16QAM o 64 QAM. Suponiendo que utilizamos la modulación de mayor eficiencia espectral, es decir 64 QAM que transmite 6 bits/símbolo, resulta que en un PRB podemos llegar a transmitir hasta 504 bits, incluyendo datos codificados de usuario, canales de control y señalización propia de la capa física del sistema, cada 0,5 ms, lo que proporciona una velocidad bruta de transmisión de pico de aproximadamente $R_{b,PRB} = 504 \text{ bits} / 0,5 \text{ ms} \approx 1 \text{ Mb/s}$. Por consiguiente, tomando en consideración el número de radiobloques indicados en la Tabla 5-2, las velocidades de pico globales en la interfaz aire para las distintas canalizaciones son:

Tabla 5-3 Velocidades de pico posibles en función de la canalización

Canalización	1,4 MHz	3 MHz	5 MHz	10 MHz	15 MHz	20 MHz.
Velocidad de Pico Total (Mb/s)	$\cong 6$	$\cong 15$	$\cong 25$	$\cong 50$	$\cong 75$	$\cong 100$
Velocidad de pico bruta de usuario (Mb/s) (15% de los recursos destinados a control y señalización)	$\cong 5,1$	$\cong 12,8$	$\cong 21$	$\cong 42,5$	$\cong 63,7$	$\cong 85$

Como se ha comentado anteriormente, los valores de velocidad de pico indicados en la primera fila incluyen tanto los datos de usuario como los canales de señalización del sistema y las señales de referencia. Si estimamos que los recursos destinados a control y señalización en LTE oscilan alrededor de un 15% aproximadamente del total disponible, entonces en la segunda fila de la Tabla 5-3 se indica una cota a las velocidades de transmisión de pico brutas o a nivel de capa física por usuario del sistema. Estas velocidades se corresponden a un sistema sin multiplexado espacial (sin MIMO). En el caso de considerar una estructura MIMO 2x2 podemos estimar grosso modo que las velocidades de pico serán el doble, es decir, se confirma que la interfaz aire del sistema LTE puede alcanzar los 150Mb/s en el enlace descendente si dispone de una canalización 20 MHz, tal como se ha indicado anteriormente. Obviamente, el cálculo anterior es una primera estimación de las capacidades de transmisión del sistema. Las velocidades de pico reales alcanzadas por un usuario dependen de muchos factores, por ejemplo la calidad del canal radio, el número de usuarios simultáneos en la celda, el tipo de despliegue realizado por el operador, el tipo de servicio considerado y la calidad (QoS) asociada, la capacidad del terminal móvil etc. No obstante, y con objeto de tener una primera estimación sobre las velocidades de pico máximas alcanzables por un terminal véase sección 5.7 dedicada a las características del terminal móvil.

5.5.2 Estructura de trama

En el dominio temporal los recursos físicos del sistema LTE se estructuran siguiendo dos posibles estructuras de trama, [9], que se describen a continuación:

Estructura de trama tipo 1

Esta estructura es válida para sistemas que utilizan duplexado por división de frecuencia (FDD) y aplica tanto al enlace descendente como al ascendente. Además, la estructura soporta tanto half como full duplex FDD.

En esta estructura el eje temporal se divide en tramas de 10 ms. Cada trama a su vez está compuesta por 20 ranuras temporales (*Slots* o TS) de duración 0,5 ms. Se define una unidad básica de recursos, formada por dos ranuras temporales (TS) denominada subtrama de duración 1 ms. Esta estructura está optimizada para coexistir con el sistema UMTS⁹.

En cada ranura temporal (TS) se transmiten 6 ó 7 símbolos OFDM cada uno de ellos de duración $T_s = 66,7 \mu s$. En el caso de utilizar 7 símbolos, el Prefijo Cíclico (CP), que a partir de ahora denominaremos prefijo cíclico normal, tiene una duración de $4,7 \mu s$, salvo para el primer símbolo que tiene un prefijo cíclico de $5,2 \mu s$. Si se utilizan 6 símbolos por ranura temporal entonces el prefijo cíclico, que a partir de ahora denominaremos prefijo cíclico largo, tiene una duración de $16,67 \mu s$. Normalmente en celdas muy grandes se utiliza una estructura de 6 símbolos por ranura temporal ya que los retardos de propagación pueden llegar a ser de algunos μs , lo que requiere un mayor prefijo cíclico para compensar la propagación multicamino (véase sección 4.1.2).

Los usuarios se ubican en determinadas subportadoras, siguiendo la estructura de un PRB, por un intervalo de tiempo preestablecido, que se determina de forma dinámica por el gestor de recursos radio (*scheduler*). La mínima granularidad temporal en la asignación de recursos es de 1 ms, es decir una subtrama.

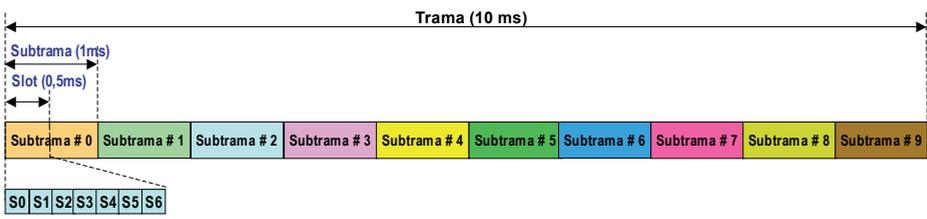


Figura 5-3 Estructura de trama tipo 1

Existe un modo de operación denominado *Reduced Subcarrier Spacing* en donde la separación entre portadoras (Δf) es de 7,5KHz. En esta modalidad se transmiten 3 símbolos OFDMA por ranura temporal utilizando un Prefijo cíclico de $33,33 \mu s$. Esta modalidad se utiliza cuando se opera en modo MBSFN (*Multimedia Broadcast Single Frequency Network*).

⁹ En el sistema UMTS la duración de una trama también es 10ms.

Estructura de trama tipo 2

Esta es una estructura pensada para operar en modo TDD (Duplexado por división en tiempo). Como en el caso anterior, el eje temporal se divide en tramas de 10 ms. Cada trama a su vez está compuesta por 10 subtramas, cada una de duración 1 ms. Es una estructura de trama mucho más flexible que la anterior en tanto que contiene subtramas de transmisión tanto para el enlace descendente (DL) como ascendente (UL) así como subtramas especiales que contienen los símbolos piloto de los enlaces ascendente y descendente y periodos de guarda entre transmisión y recepción. Esta subtrama especial aparece para facilitar la transición entre los enlaces descendente y ascendente por razones de sincronización. En efecto, aunque el sistema esté correctamente sincronizado temporalmente, por lo que respecta al enlace ascendente sus transmisiones siempre pueden sufrir de una pequeña incertidumbre temporal que obliga a mantener un mínimo tiempo de guarda.

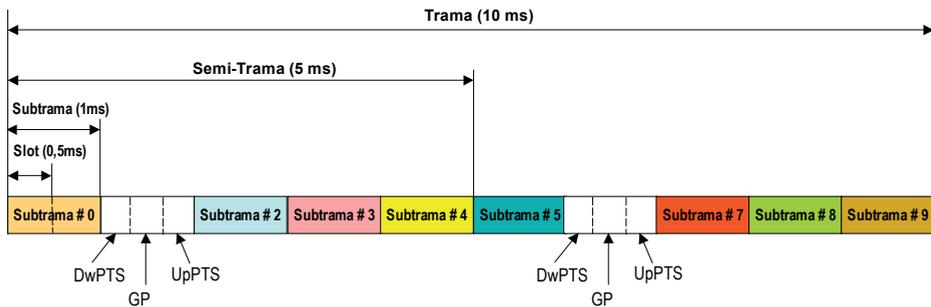


Figura 5-4 Estructura de trama tipo 2

La subtrama especial tiene una configuración variable. En ella se pueden distinguir tres campos:

- DwPTS, que corresponde a la transmisión en el enlace descendente, y cuya longitud mínima es 1 símbolo OFDM. Este primer símbolo OFDM se utiliza para ubicar, ahora que el sistema opera en modo TDD, la denominada señal de sincronización primaria. Los restantes símbolos de este campo (si existen) pueden transportar señales de referencia o incluso datos.
- UpPTS que corresponde a la transmisión en el enlace ascendente. Se utiliza para transmitir un preámbulo corto de acceso aleatorio (*short RACH preamble*) compuesto por 2 símbolos OFDM¹⁰. Los restantes símbolos de este campo (si existen) pueden transportar señales de referencia o incluso datos.
- GP o periodo de guarda. La longitud del mismo depende de los campos anteriores.

Hay 7 combinaciones diferentes de ubicación de las distintas subtramas, tal como muestra la siguiente tabla, si bien las subtramas 0 y 5 son siempre para la transmisión en el enlace descendente y la subtrama 1 es siempre de tipo especial.

¹⁰ Hay que destacar que en el caso de utilizar preámbulos de acceso aleatorio largos, éstos se transmiten utilizando subtramas convencionales de 1 ms.

Tabla 5-4 Posibles configuraciones de las subtramas. Sistema operando en modo TDD

Configuración	Periodicidad de la subtrama específica	Número de Subtrama									
		0	1	2	3	4	5	6	7	8	9
0	5 ms	D	S	U	U	U	D	S	U	U	U
1	5 ms	D	S	U	U	D	D	S	U	U	D
2	5 ms	D	S	U	D	D	D	S	U	D	D
3	10 ms	D	S	U	U	U	D	D	D	D	D
4	10 ms	D	S	U	U	D	D	D	D	D	D
5	10 ms	D	S	U	D	D	D	D	D	D	D
6	5 ms	D	S	U	U	U	D	S	U	U	D

En la tabla “D” indica subtrama dedicada al enlace descendente, “U” subtrama dedicada al enlace ascendente y “S” subtrama especial.

Como se ha dicho anteriormente, las tramas 0 y 5 siempre corresponden al enlace descendente ya que incluyen a las señales de sincronización utilizadas para la búsqueda inicial de celda o de celdas adyacentes.

5.5.3 Enlace Descendente

En este apartado se describen las principales características de las señales físicas que permiten al terminal móvil sincronizarse al sistema y demodular coherentemente las señales OFDMA transmitidas en el enlace descendente. Además, también se revisan las principales características de los procedimientos físicos (canales físicos) utilizados para transportar la información destinada a los usuarios situados en la zona de cobertura del correspondiente eNB, así como de los mecanismos de señalización y control que posibilitan una transmisión eficiente y controlada de la información. El apartado concluye indicando cómo los diferentes canales definidos están ubicados en los recursos frecuencia-tiempo disponibles.

5.5.3.1 Señales Físicas en el enlace descendente:

Con el fin de facilitar la implementación de los mecanismos físicos necesarios para realizar la demodulación/detección de las señales moduladas OFDMA propias del enlace descendente del sistema LTE, se han establecido un conjunto de señales físicas de referencia y de sincronización temporal, que se describen a continuación.

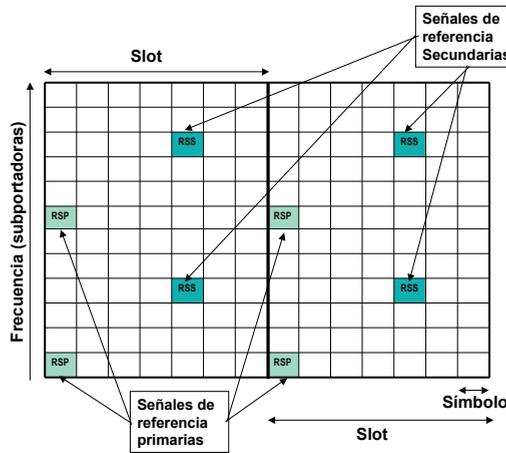
Señales de Referencia (RS)

Las señales de referencia o símbolos piloto se utilizan para:

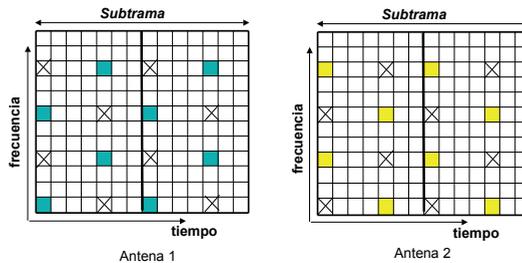
- Obtener medidas de calidad en el enlace descendente.

- Estimar la respuesta impulsional del canal¹¹ para demodulación/detección coherente.
- Implementar mecanismos de búsqueda de celda y sincronización inicial.

Las señales de referencia se sitúan en determinados recursos elementales (una subportadora y un símbolo) y mantienen una separación en el dominio de la frecuencia de seis subportadoras. Como mínimo se necesitan dos señales de referencia por PRB denominadas señales de referencia primaria (RSP). Estas señales de referencia primarias están ubicadas en el primer símbolo de una ranura temporal (TS). Si las condiciones del canal radio lo requieren puede ubicarse un segundo grupo de señales piloto o de referencia, denominadas señales de referencia secundarias (RSS), típicamente situadas en el quinto símbolo de una ranura temporal (TS). Véase Figura 5-5 (A) en donde se muestra un ejemplo típico de la ubicación de las señales de referencia primarias y secundarias dentro de una subtrama genérica, compuesta por dos bloques de recursos físicos (PRBs).



(A) Caso de una única antena. El ejemplo muestra las señales de referencia para una subtrama con 12 subportadoras



(B) Caso de multiplexado espacial con dos antenas. Como en el caso anterior se muestra un ejemplo con 12 subportadoras para una subtrama. Los recursos frecuencia-tiempo marcados con "X" indican que no se transmite en ellos.

Figura 5-5 Ejemplos de ubicación de las señales de referencia

¹¹ El estándar LTE no establece el procedimiento concreto a seguir para estimar la respuesta impulsional del canal, ya que esto depende de la implementación concreta de cada fabricante. En la literatura se han publicado diversos trabajos sobre el tema. Véase por ejemplo, referencia [15].

En la definición de los símbolos de referencia primarios y secundarios se ha considerado:

- La separación temporal de los símbolos de referencia depende de la máxima frecuencia Doppler (f_d) del canal que el sistema puede soportar. En efecto, entendiendo la frecuencia Doppler como la máxima velocidad con que el canal puede variar, el tiempo de coherencia del canal (es decir el tiempo en el que el canal no varía) es aproximadamente $T_c = 1/f_d$, y, tomando en consideración el criterio de Nyquist, para poder reconstruir la respuesta impulsional a partir de los símbolos de referencia será necesario ubicar al menos 2 símbolos por T_c . En el caso del sistema LTE la frecuencia Doppler máxima es de aproximadamente 950 Hz (correspondiente a una máxima velocidad de 500Km/h y una frecuencia de operación de 2 GHz), por consiguiente $T_c \cong 1\text{ms}$ y por lo tanto la separación mínima entre símbolos de referencia debe ser 0,5 ms. Es decir, al menos hay que ubicar un símbolo de referencia por slot, tal como se plantea para los símbolos de referencia primarios. Cuando el sistema utiliza símbolos de referencia secundarios (lo que es habitual), se aumenta la precisión de la estima del canal ya que se ubican 4 símbolos de referencia por T_c .
- En la dimensión frecuencial, el sistema ubica un símbolo de referencia cada seis subportadoras. Nuevamente esta separación está originada por el comportamiento del canal móvil. En entornos urbanos, el valor de la dispersión temporal del canal (medida en términos del *Delay Spread* o σ_s) es del orden de $1\mu\text{s}$. Por lo tanto la banda de coherencia del canal puede estimarse en $B_c \cong 1/10\sigma_s$, es decir del orden de los 100KHz. Por lo tanto, para que sea posible estimar por interpolación, y a partir de las señales de referencia, la función de transferencia del canal será necesario que la separación de dichos símbolos sea inferior a la banda de coherencia del canal, tal como ocurre en el sistema LTE en donde la separación frecuencia entre símbolos de referencia es 6 subportadoras es decir 90 KHz. Si además de los símbolos de referencia primarios, consideramos también los secundarios, entonces la separación entre subportadoras que ubican símbolos de referencia es de sólo tres subportadoras, es decir 45 KHz,. Nuevamente la presencia de símbolos de referencia secundarios sirve para aumentar la precisión con que se estima el canal (ahora en el dominio de la frecuencia) .

En el caso de múltiples antenas en el transmisor entonces cada puerto (antena) tiene su correspondiente patrón de señales piloto, tal como se muestra en la Figura 5-5 (B).

Finalmente, indicar que el estándar LTE define un total de 504 secuencias de señales de referencia definidas a partir de secuencias de Gold¹² de longitud 31. Estas 504 secuencias distintas están agrupadas en 168 grupos de tres secuencias, donde cada grupo corresponde a una identidad de un emplazamiento (eNB) distinto. Hay que subrayar que se necesitan 504 secuencias diferentes ya que está permitida la sectorización de los emplazamientos. En efecto, típicamente hay tres sectores por eNB por lo que: 168 identidades distintas de eNB x 3 sectores por emplazamiento precisan de 504 secuencias distintas para identificar las celdas.

¹² Véase anexo 5.1 de este capítulo para detalles sobre las propiedades de las secuencias de Gold.

Señal de Sincronización (SCH)

La señal de sincronización (SCH) se utiliza para facilitar los procesos de sincronización temporal (fundamentalmente a nivel de trama y subtrama) del sistema. Esta señal se desglosa en:

- P-SCH (*Primary SCH*) que permite la sincronización temporal a nivel de subtrama, mediante procedimientos de correlación entre la señal recibida y una secuencia de referencia almacenada en el receptor.
- S-SCH (*Secondary SCH*) que posibilita la sincronización temporal a nivel de trama, utilizando la misma metodología de correlación temporal.

La señal P-SCH se genera a partir de una secuencia de Zadoff-Chu¹³, [12], [13] de longitud 63 a la que se le ha eliminado el elemento central (por lo tanto sólo se transmiten 62 símbolos de dicha secuencia). Se transmite utilizando las 62 subportadoras centrales de la banda utilizada¹⁴ por el sistema, dejando sin ocupar la subportadora central o portadora DC. Además, y para separar esta señal de otras señales transmitidas utilizando el resto de subportadoras, se dejan 5 subportadoras vacías a su izquierda y derecha, que actúan como bandas de guarda. Es por ello que en algunos casos dice que la señal P-SCH ocupa 72 subportadoras. La secuencia elegida se caracteriza por tener muy buenas propiedades de correlación y ser muy robusta frente a derivas (*offsets*) en frecuencia, lo que facilita la sincronización temporal inicial del sistema.

La señal S-SCH siempre está asociada a una señal P-SCH. Utiliza secuencias de máxima longitud o *m-sequences*¹⁵. Cada secuencia secundaria de sincronización se obtiene como resultado de entrelazar, en el dominio de la frecuencia, una secuencias de máxima longitud de longitud 31 y su correspondiente versión desplazada cíclicamente (véase Figura 5-6). Como en el caso de la señal de sincronización primaria, la señal S-SCH también se transmite utilizando 62 subportadoras centrales de la banda disponible y también se dejan 5 subportadoras vacías a izquierda y derecha para separarla del resto de señales transportadas por otras subportadoras. La modulación utilizada en cada subportadora es BPSK.

Además para facilitar la sincronización temporal a nivel de trama, la señal secundaria se transmite dos veces por trama, en ranuras temporales (TS) convenientemente espaciadas. La secuencia de longitud 62 utilizada en cada símbolo no es exactamente la misma ya que difiere en la forma en que se genera. En efecto, en el caso de la secuencia asignada al segundo símbolo, los papeles de las secuencias de máxima longitud de la Figura 5-6 están intercambiados, es decir, los símbolos de la secuencia de máxima longitud original ocupan ahora las posiciones pares y los símbolos asociados a la secuencia rotada cíclicamente las posiciones impares de la secuencia de 62 símbolos.

¹³ Las secuencias Zadoff-Chu pertenecen a la familia de secuencias de CAZAC (Constant Amplitude Zero Correlation Codes). Véase Anexo 5.1 para una descripción de las características y propiedades de dichas secuencias.

¹⁴ Recuérdese que el sistema LTE permite anchos de banda de 1.4, 3, 5, 10, 15 ó 20 MHz.

¹⁵ Véase sección I-3 del Anexo 5.1 de este capítulo.

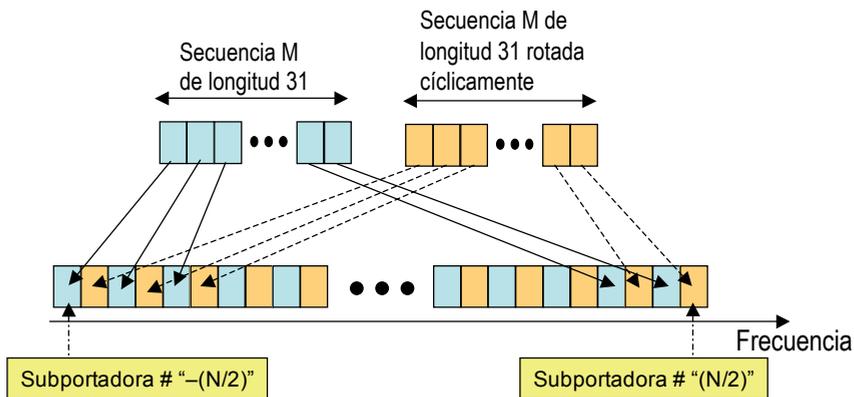


Figura 5-6 Secuencia de Sincronización Secundaria

Como en el caso de la secuencia de sincronización primaria, la secuencia de sincronización secundaria así construida se caracteriza por tener muy buenas propiedades de correlación y ser muy robusta frente a derivas (*offsets*) en frecuencia.

Finalmente, indicar que la señal S-SCH define la identidad, entre 168 posibles valores, de la celda a la que se sincroniza el terminal móvil. En efecto, como al conectar (activar) el terminal móvil el primer procedimiento que éste realiza es el procedimiento de sincronización y adquisición inicial de parámetros (véase sección 5.8.1 de éste capítulo), las señales de sincronización primaria y secundaria se detectan en primer lugar, de modo que la identificación de celda indicada por éstas señales se utiliza para posteriormente identificar en qué recursos (posición en la estructura frecuencia-tiempo) se van a situar las señales de referencia primarias (RSP) y secundarias (RSS).

Hay que destacar que, al situar ambas señales de sincronización en el centro del radiocanal disponible, se facilita la sincronización de los terminales móviles a la red, independientemente de cual sea el número de subportadoras con las que dichos terminales móviles operen. O dicho de otra forma, el mecanismo de sincronización temporal es siempre el mismo, independiente de la canalización del sistema. Recuérdese que el sistema permite una canalización flexible entre 1.4 y 20 MHz.

En el caso de utilizar una estructura de trama del tipo 1 (denominada en el estándar LTE como SF1) la señal P-SCH se transmite utilizando el último símbolo de las ranuras temporales (TS) 0 y 10 (correspondientes a la primera y sexta subtrama), mientras que la señal S-SCH se transmite utilizando en el penúltimo símbolo de las mismas ranuras temporales.

Si se utiliza una estructura de trama del tipo 2 (denominada SF2 en el estándar), entonces la señal P-SCH se transmite utilizando el campo DwPTS de la subtrama especial, mientras que la señal S-SCH se transmite utilizando el último símbolo OFDM de las subtramas numeradas como cero y cinco de dicha estructura de trama.

5.5.3.2 Canales Físicos en el enlace descendente

En este apartado se describen brevemente las características más relevantes de los canales físicos especificados para el enlace descendente en el sistema LTE. Se diferencia entre canales físicos de tráfico y canales físicos de control.

A diferencia de sistemas de comunicaciones celulares anteriores como GSM o UMTS, el sistema LTE no contempla el uso de canales de tráfico dedicados, sino compartidos. Es decir, sólo se establecen cuando un usuario (terminal móvil) debe recibir datos destinados a él. Existe la posibilidad de establecer enlaces punto a punto (entre un terminal móvil y la red) o punto multipunto (entre la red y múltiples usuarios, es decir una conexión tipo *multicast*).

Para permitir una transmisión eficiente de los datos de los usuarios a través de la interfaz aire, también se necesitan los denominados canales físicos de control. Estos canales transportan tanto señales de control generadas a nivel de capa física como aquel tipo de mensajes de control generados en las capas superiores del sistema.

CANALES FÍSICOS DE TRÁFICO

Los canales físicos de tráfico enlace descendente, se clasifican en:

- *Physical Downlink Shared Channel* (PDSCH).
- *Physical Multicast Channel* (PMCH).

Physical Downlink Shared Channel (PDSCH)

El canal PDSCH transmite habitualmente información de usuario. Contiene la información entregada por la capa MAC mediante el canal de transporte *Downlink Shared Channel* (DL-SCH). También puede transportar información de aviso (PCH) y aquella de radiodifusión que no sea imprescindible para que el terminal móvil se enganche a la red. En particular, transporta los denominados *System Information Blocks* o SIBs (véase anexo 5.2).

No es un canal dedicado, sólo se asigna al usuario cuando éste tiene algo que recibir en el enlace descendente (DL) y como resultado de la gestión de recursos realizada por el algoritmo de *scheduling* de dicho enlace.

Los posibles esquemas de modulación considerados son: QPSK, 16QAM y 64 QAM. En general se utilizan turbo códigos de tasa 1/3 para la codificación de canal. El canal PDSCH permite utilizar diferentes tipos de redundancia para implementar el mecanismo de retransmisión híbrido (HARQ). La granularidad temporal (*Transmission Time Interval* o TTI) asociada a este canal es 1 ms, es decir una subtrama.

El canal PDSCH se mapea en el dominio frecuencia-tiempo utilizando los denominados *Physical Resource Blocks* (PRB).

Existen siete modos de operación de este canal, tal como se indica en la siguiente tabla:

Tabla 5-5.- Modos de operación del canal PDSCH

Operación	Características Principales
Modo 1	Transmisión por una única antena en el eNB
Modo 2	Diversidad en transmisión
Modo 3	Multiplexado espacial en bucle abierto
Modo 4	Multiplexado espacial en bucle cerrado
Modo 5	MIMO para esquemas multiusuario
Modo 6	Conformación de haz con realimentación desde el terminal móvil para indicar qué tipo de pre-codificación debe utilizarse en el proceso de conformación de haz.
Modo 7	Conformación de haz sin realimentación. En este caso el eNB determina el esquema de pre-codificación apropiado para la conformación de haz a partir, por ejemplo de los ángulos de llegada de canales del enlace ascendente.

Nota.- Véase capítulo 4 para detalles sobre los posibles métodos de transmisión utilizando técnicas MIMO.

Physical Multicast Channel (PMCH)

El *Physical Multicast Channel* (PMCH) se utiliza para transportar información MBMS (*Multimedia Broadcast and Multicast System*) en redes iso-frecuencia (*Single-Frequency Networks* o *SFN*). Véase anexo 5.4 para más detalles sobre este tipo de redes y servicios.

La estructura básica del canal PMCH es muy similar a la del canal PDSCH. Las principales diferencias son:

- La disposición de los símbolos de referencia en el canal PMCH difiere de la utilizada en el canal PDSCH. De hecho en una misma subtrama no pueden coexistir los canales PMCH y PDSCH o lo que es lo mismo, deben existir subtramas específicas sólo dedicadas a transmitir el canal PMCH cuando el sistema permite usuarios *multicast*.
- Se utiliza siempre el prefijo cíclico extendido, ya que ahora el terminal móvil puede recibir la señal de múltiples eNBs, por lo que los retardos por propagación multicamino (asociados a los diversos eNB) aumentan. Véase capítulo 4 para detalles sobre la transmisión en redes iso-frecuencia.
- Puesto que la señal transmitida es observada por múltiples terminales móviles de forma simultánea, no es posible aplicar el mecanismo de HARQ.
- Tampoco es posible gestionar de forma dinámica en cada eNB, ni los formatos de trans-

misión ni los recursos asociados, ya que todos los eNBs que forman parte de la red iso-frecuencia deben enviar la misma señal y de forma sincronizada.

- Los canales de señalización o control PDCCH y PHICH (que se describirán posteriormente en el sub-apartado destinado a canales de control) no pueden ocupar más de dos símbolos OFDM en la subtrama MBSFN¹⁶ (*MBMS Single Frequency Network*). En este tipo de subtrama el canal PDCCH se utiliza únicamente para transmitir las asignaciones de recursos en el enlace ascendente, pero no soporta ningún tipo de señalización correspondiente al canal PMCH, ya que la asignación de recursos para los datos del servicio MBMS se efectúa utilizando señalización de capas superiores.

CANALES DE CONTROL

Por lo que respecta a los canales físicos de control del enlace descendente, éstos son:

- *Physical Broadcast Channel* (PBCH).
- *Physical Downlink Control Channel* (PDCCH).
- *Physical Control Format Indicator Channel* (PCFICH).
- *Physical Hybrid ARQ Indicator Channel* (PHICH).

Physical Broadcast Channel (PBCH)

Este canal transporta información básica sobre la red, la denominada *Master Information Block* (MIB)¹⁷. Esta información consiste en:

- 4 bits para identificar la canalización utilizada en la celda. Ello es necesario ya que el sistema es flexible y permite diferentes anchos de banda desde 1.4 MHz hasta 20 MHz.
- 3 bits para definir la estructura del canal PHICH que se utiliza para transportar información de reconocimientos del mecanismo de retransmisión híbrida HARQ.
- 7 bits correspondientes al denominado número de identificación de trama (*System Frame Number-SFN*).

Estos 14 bits junto con 10 bits de relleno ("*spare bits*") conforman un bloque de transporte básico al que se añade un código de redundancia cíclico de 16 bits para detectar errores en recepción. Este bloque básico de 40 bits pasa a través de un código convolucional de tasa 1/3 para dar lugar a 120 bits codificados. Estos 120 bits codificados se repiten de forma sistemática 16 veces hasta dar lugar a los 1920 bits que se transmiten en el canal PBCH, cuando se considera un prefijo cíclico (CP) corto (4,7 ó 5,2 μ s). En el caso de considerar un prefijo cíclico

¹⁶ Véase el anexo 5.4 de este capítulo para una mayor información sobre la estructura de la subtrama MBSFN.

¹⁷ Véase anexo 5.2 de este capítulo.

largo (16,67 μ s), el procedimiento es idéntico sólo que ahora se utilizan 6 bits de relleno lo que da lugar a la transmisión de 1728 bit codificados. Por consiguiente, puede considerarse que la tasa efectiva de codificación es de 1/48, por lo que éste es un canal con una elevada protección lo que garantiza su recepción correcta incluso cuando el terminal móvil se encuentra en el límite de la zona de cobertura de la celda.

El canal PBCH utiliza modulación QPSK y aleatorización mediante una secuencia que es específica de cada celda. El canal PBCH se transmite en 4 tramas consecutivas (es decir con un TTI=40ms) en las ranuras temporales 0,1 2 y 3 de la primera subtrama utilizando 72 subportadoras centradas alrededor de la subportadora central o subportadora DC y descontando los símbolos piloto. Este canal no necesita señalización explícita para ser demodulado y se puede decodificar a partir de una sola recepción.

Physical Downlink Control Channel (PDCCH)

El canal PDCCH transporta los denominados *Downlink Control Information* (DCI), que incluyen información de asignación de recursos para los canales de aviso (PCH) y de tráfico compartido (DL-SCH). También transporta información sobre los mecanismos de retransmisión (*Hybrid ARQ*)¹⁸.

Cada DCI está formado por la agregación uno o varios *Control Channel Elements* (CCE). Cada CCE corresponde a 9 conjuntos de cuatro parámetros o informaciones (denominados *Resource Element Groups* o REGs) relacionadas con la capa física. Por ejemplo, en un CCE se puede transmitir el esquema de modulación-codificación utilizado, información sobre el ajuste del control de potencia del enlace ascendente, información sobre el mecanismo de retransmisión híbrida (HARQ) y la versión de redundancia aplicada, etc. Cada REG se mapea mediante 4 símbolos QPSK, es decir ocupa 4 subportadoras, formando los denominados mini CCE.

Por consiguiente, un CCE ocupa 36 subportadoras a las que debe añadirse, si es el caso, el número de subportadoras ocupadas por las señales de referencia. El número de CCEs asignados a un canal PDCCH lo determina cada eNB en función de las condiciones del canal, ya que hay una relación implícita entre las condiciones del canal, el número de retransmisiones efectuadas en el enlace ascendente y los recursos de señalización y control necesarios para transmitir los reconocimientos ACK/NACK en el enlace descendente. Por ejemplo, en el caso de terminales móviles ubicados cerca del eNB probablemente con un único CCE sea suficiente ya que las condiciones de propagación serán habitualmente buenas y el número de retransmisiones (si las hay) bajo. Sin embargo, si el terminal móvil está lejos del eNB, y las condiciones del canal empeoran, entonces se necesitará un número mayor de CCEs para conseguir suficiente robustez en la transmisión.

¹⁸ Véase anexo 5.3 de este capítulo para detalles sobre el mecanismo de retransmisión híbrida y los mecanismos de redundancia utilizados en LTE.

El canal PDCCH se ubica en los “n” primeros símbolos OFDMA de cada sub-trama ($n \leq 3$) y existen cuatro posibles formatos:

Tabla 5-6 Configuraciones del canal PDCCH

Formato PDCCH	Número de CCEs	Número de Regs o mini-CCEs	Número de bits del PDCCH
0	1	9	72
1	2	18	144
2	4	36	288
3	8	72	576

Obsérvese que como se utiliza modulación QPSK (transmite 2 bits por símbolo) se necesitan 36 subportadoras por CCE y como estas subportadoras se agrupan de cuatro en cuatro existe un total de 9 mini-CCEs por CCE, tal como se muestra en la Tabla 5-6.

El tipo de información que transportan los DCIs, está indicado en la Tabla 5-7

Tabla 5-7 Formatos DCI

Formato	Propósito	Información Transmitida
0	Reconocimiento petición asignación recursos para PUSCH (ver sección 5.5.4)	Transmite la asignación de recursos para un PUSCH. Incluye: <ul style="list-style-type: none"> Asignación bloque de recursos (PRBs). Esquema de modulación-codificación. Información sobre HARQ y versión de redundancia. Comandos del control de potencia a aplicar al PUSCH. Petición de transmisión de información del estado del canal (<i>CQI request</i>) en modo no periódico.
1	Asignación para PDSCH (modo simple)	Transmite la asignación de recursos para un PDSCH operando en los modos 1,2 y 7 (ver Tabla 5-5). Incluye: <ul style="list-style-type: none"> Asignación bloque de recursos (PRBs). Esquema de modulación-codificación. Información sobre HARQ. Comandos del control de potencia a aplicar al canal PUCCH (<i>Physical Uplink Control Channel</i>) definido en sección 5.5.4.1.
1A	Asignación para PDSCH en formato compacto	Transmite la asignación (de forma compactada) de recursos para un PDSCH operando en los modos 1,2 y 7 y asignación de recursos para acceso aleatorio regulado ¹⁹ . Incluye: <ul style="list-style-type: none"> Asignación bloque de recursos (PRBs) para PDSCH o PRACH (canal de acceso aleatorio). Esquema de modulación-codificación. Información sobre HARQ. Comandos del control de potencia a aplicar al PUCCH.
1B	Asignación para PDSCH operando con conformación de haz	Transmite la asignación (de forma compactada) de recursos para un PDSCH operando con conformación de haz mediante realimentación desde el terminal móvil (modo 6 de operación del PDSCH). La información transmitida es la misma que en el formato 1A pero ahora se incluye un indicador del tipo de pre-codificación que debe utilizarse en el proceso de conformación de haz.

¹⁹ Véase sección 5.8.2 sobre mecanismos de acceso aleatorio.

1C	Asignación para PDSCH en formato muy compacto	Transmite la asignación (en forma compactada) de recursos para un PDSCH. Cuando se utiliza el formato 1C el canal PDSCH sólo puede utilizar modulación QPSK. Este modo de operación es utilizado habitualmente para transmitir mensajes de aviso o mensajes de radiodifusión e incluye: <ul style="list-style-type: none"> • Asignación bloque de recursos (PRBs). • Esquema de modulación-codificación. • Versión de redundancia.
1D	Asignación para PDSCH caso MIMO multiusuario	Transmite la asignación (en forma compactada) de recursos para un PDSCH cuando el eNB considera técnicas MIMO para esquemas multiusuario. La información transmitida es la misma que la correspondiente al formato 1B, pero en lugar de incluir información sobre tipo de pre-codificación ahora se incluye información para indicar si la potencia de transmisión es compartida (ó no) por múltiples usuarios.
2	Asignación para PDSCH con MIMO operando en bucle cerrado	Transmite la asignación de recursos para un PDSCH operando en el modo 4 (MIMO en bucle cerrado). Incluye: <ul style="list-style-type: none"> • Asignación bloque de recursos (PRBs). • Información sobre HARQ. • Comandos del control de potencia a aplicar al PUCCH. • Número de caminos espaciales. • Esquema de modulación-codificación para cada camino. • Información de precodificación.
2A	Asignación para PDSCH con MIMO operando en bucle abierto	Transmite la asignación de recursos para un PDSCH operando en el modo 3 (MIMO en bucle abierto). La información transmitida es la misma que para el modo 2 excepto que si sólo hay dos antenas se transmite información de precodificación. Si hay cuatro antenas en el eNB entonces se transmiten dos bits que indican el rango del MIMO.
3	Control de potencia para PUCCH ²⁰ y PUSCH. Ajuste de 2 bits	Transmite comandos de ajuste del control de potencia para un grupo de terminales móviles. El ajuste se indica mediante 2 bits por cada terminal móvil.
3A	Control de potencia para PUCCH y PUSCH. Ajuste de 1 bit	Transmite comandos de ajuste del control de potencia para un grupo de terminales. El ajuste se indica mediante 1 bit por cada terminal móvil.

Physical Control Format Indicator Channel (PCFICH)

Informa al terminal móvil sobre el número de símbolos (1,2,3) utilizados para transmitir el canal PDCCH (1,2 ó 3), utilizando un indicador denominado *Channel Format Indicator* o (CFI). Se transmite en cada subtrama, siempre en el primer símbolo. EL CFI está formado por 32 bits (16 símbolos) que se transmiten utilizando una modulación QPSK.

²⁰ Los canales PUCCH y PUSCH están definidos en la sección 5.5.4.1.

El canal PCFICH se mapea en el primer símbolo OFDM de cada subtrama en cuatro grupos de cuatro subportadoras repartidas uniformemente. La posición del primer grupo de subportadoras viene determinada por el valor del identificador de celda (Cell-ID) y de la canalización (ancho de banda) del sistema. Esta dependencia permite que en recepción se conozcan a priori las subportadoras en donde se mapea el canal PCFICH, ya que al sincronizarse temporalmente se habrá determinado, a partir del canal de sincronización secundario (S-SCH), la identidad de la celda (Cell-ID) y se conocerá, a partir del canal de radiodifusión (PBCH), el ancho de banda o canalización utilizada por el eNB.

Cuando se utilizan técnicas MIMO el canal PCFICH se transmite por la misma antena (*port*) que la señal PBCH.

En el caso de portadoras dedicadas únicamente a la transmisión en modo MBSFN, este canal no se transmite ya que la subtrama MBMS no contiene información de control. (véase anexo 5.4 de este capítulo).

Physical Hybrid ARQ Indicator Channel (PHICH)

Transporta los reconocimientos (ACK/NAK) para implementar el mecanismo de HARQ en el enlace ascendente. Utiliza modulación BPSK y códigos de repetición de tasa 1/3. La información de múltiples usuarios se multiplexa mediante técnicas por división en código (véase Figura 1-7) utilizando un conjunto de secuencias ortogonales de Walsh²¹ predefinidas.

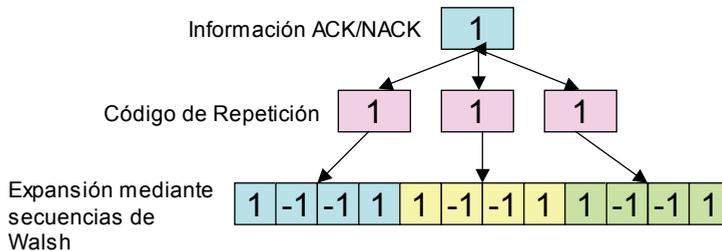


Figura 5-7.-Construcción de la señal PHICH

Un canal PHICH está formado por la información ACK/NAK (1 bit) correspondiente hasta, como mucho, 8 terminales móviles. En caso de utilizar un prefijo cíclico (CP) largo entonces un grupo PHICH contiene información correspondiente a sólo 4 terminales móviles.

Las secuencias ortogonales tienen un factor de ensanchamiento (SF) de valor 4 en el caso de CP normal y SF de 2 en el caso de CP extendido (largo) tal como se muestra en la figura 5.7 o en la tabla 5.8.

²¹ Véase anexo 5.1 para la descripción de las propiedades ortogonalidad de las secuencias de Walsh.

Tabla 5-8 Configuraciones canal PHICH

Índice de Secuencia	Secuencias Ortogonales	
	CP Normal (SF=4)	CP extendido (SF=2)
0	+1,+1,+1,+1	+1,+1
1	+1,-1,+1,-1	+1,-1
2	+1,+1,-1,-1	+j,+j
3	+1,-1,-1,+1	+j,-j
4	+j,+j,+j,+j	---
5	+j,-j,+j,-j	---
6	+j,+j,-j,-j	---
7	+j,-j,+j,-j	---

En la Tabla 5-8 se muestran las posibles configuraciones utilizadas en el sistema. En el caso de prefijo cíclico normal, (CP normal) hay ocho posibles secuencias ortogonales, cada una de ellas destinada a uno de los ocho posibles terminales móviles. Las cuatro primeras secuencias, son ortogonales entre ellas y se transmiten moduladas BPSK utilizando la componente en fase (I), mientras que las cuatro últimas también son ortogonales entre ellas y se transmiten utilizando la componente en cuadratura (Q). En el caso de prefijo cíclico largo o extendido, el proceso es el mismo, sólo que ahora las secuencias ortogonales están compuestas por sólo dos bits, y hay cuatro posibles secuencias. Análogamente las dos primeras (que son ortogonales entre si) se transmiten moduladas BPSK utilizando la componente en fase (I) y las dos restantes (también ortogonales entre sí) se transmiten utilizando la componente en cuadratura (Q).

El canal de radiodifusión físico (PBCH) señala el número de grupos PHICH y su ubicación temporal, mediante 2 bits.

El canal PHICH se mapea en los símbolos OFDM de control de cada subtrama en cuatro grupos de cuatro subportadoras repartidas uniformemente. De forma análoga al caso del canal PCFICH, la posición del primer grupo de subportadoras depende del ancho de banda del sistema (número de subportadoras) y del valor del identificador de celda (Cell_ID).

Mapeo de los canales físicos sobre PRBs

Antes de proceder al mapeo de los diferentes canales físicos utilizados en el enlace descendente, y a modo de recordatorio, en la Tabla 5-9 se describen dichos canales.

Tabla 5-9 Resumen de los canales físicos del enlace descendente

Nombre Canal	Descripción
<i>Physical Downlink Shared Channel</i> (PDSCH)	Transmite información de usuario, de control tanto dedicada como común (p.e., aviso o <i>paging</i> , y de ciertos parámetros de radiodifusión)
<i>Physical Multicast Channel</i> (PMCH)	Se utiliza para transportar información MBMS destinada a los usuarios unicast en redes isofrecuencia (SFN)
<i>Physical Broadcast Channel</i> (PBCH)	Transporta la información de radiodifusión básica de la red, que permite la conexión inicial de un terminal móvil a la misma.
<i>Physical Downlink Control Channel</i> (PDCCH)	Canal de control genérico del enlace descendente. Informa sobre los recursos en el enlace descendente asignados al PDSCH.
<i>Physical Control Format Indicator Channel</i> (PCFICH)	Informa al terminal móvil sobre el número de símbolos (1,2 ó 3) utilizados para transmitir el PDCCH.
<i>Physical Hybrid ARQ Indicator Channel</i> (PHICH)	Transporta información de reconocimientos (ACK/NACK) correspondientes a las transmisiones del enlace ascendente.

En la Figura 5-8 se indica como se ubican en la estructura frecuencia-tiempo los diferentes canales físicos del enlace descendente para el caso de FDD, mientras que en la Figura 5-9 se proporciona el detalle la ubicación del contenido del canal PDCCH.

En primer lugar las señales de sincronización primaria y secundaria, P-SCH y S-SCH respectivamente, que se transmiten mediante 62 subportadoras, que aparecen ubicadas a izquierda y derecha de la subportadora central (*DC- subcarrier*) y en los símbolos 6 y 5 respectivamente de la primera ranura temporal (TS) de las subtramas numeradas como #0 y #5. Esto se hace así ya que, al estar ubicados alrededor de la subportadora central, se garantiza que, independientemente del ancho de banda del sistema considerado (es decir, independientemente del número de subportadoras disponibles en el sistema), los terminales móviles siempre podrán detectar dichas señales de sincronización, aún cuando inicialmente desconozcan el ancho de banda del sistema al que quieren acceder. En efecto, tomando en consideración que la separación entre subportadoras es 15 KHz, la banda ocupada por las subportadoras asociadas a las señales de sincronización (incluyendo la portadora DC) es de 945 KHz. Por lo tanto incluso para la canalización más estrecha prevista 1,4 MHz es posible detectar estas señales. Es conveniente recordar aquí que el sistema tiene estandarizados diferentes anchos de banda (canalizaciones) disponibles y que durante el mecanismo de sincronización inicial el terminal móvil no tiene porqué conocer el número de subportadoras con que trabaja el sistema, información que posteriormente difundirá el canal de radiodifusión (PBCH). Adicionalmente, para ambas señales (P-SCH y S-SCH) se dejan 5 subportadoras vacías a izquierda y derecha de los canales de sincronización, actuando como banda de guarda de las mismas.

El canal de radiodifusión (PBCH) se ubica utilizando los símbolos 0, 1, y 2 de la segunda ranura temporal (Slot) de la subtrama numerada como #0. Ocupa 72 subportadoras también situadas 36 a la izquierda y 36 a la derecha de la subportadora central (*DC-subcarrier*). El canal de radiodifusión aparece inmediatamente después de las señales de sincronización, ya que es necesario conocer el contenido de este canal, que define los parámetros básicos del sistema, para posteriormente poder realizar cualquier intercambio de información/señalización con la red. También es importante subrayar que, aunque no está indicado en la figura, en estas 72 subportadoras PBCH también se ubican las señales de referencia (ó señales piloto). Obviamente en el recurso frecuencia-tiempo donde se ubican las señales de referencia no se transmite ningún bit del canal PBCH.

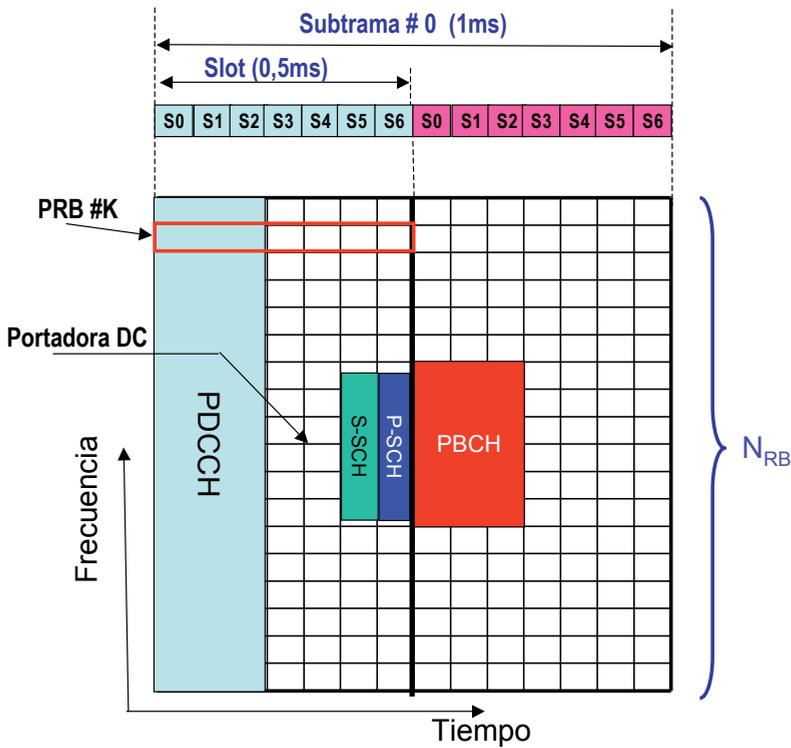


Figura 5-8 Ubicación en la estructura frecuencia-tiempo de los canales comunes de control y señales de referencia para el enlace descendente

En cada subtrama los primeros símbolos (entre 0 y 2 según la figura) de la misma los ocupa el canal PDCCH (véase la Figura 5-8 como ejemplo de ubicación de este canal). El número de símbolos OFDM utilizados para transmitir el canal PDCCH se indica mediante el canal PCFICH,

que contiene el denominado “*Control Format Indicator*” (CFI). El canal PCFICH siempre se transmite en el primer símbolo²² de la subtrama, tal como se indica en la Figura 5-9, utilizando 16 subportadoras distribuidas entre toda la banda y de acuerdo con un patrón predeterminado (véase descripción del canal (PCFICH) en la sección anterior).

Para minimizar la posibilidad de detectar por error un canal PCFICH procedente de una celda vecina, se aplica un mecanismo de desplazamiento (*offset*) en frecuencia de la posición del canal PCFICH, desplazamiento cuyo valor depende del identificador de celda (*Cell-ID*)²³. Con ello se garantiza que los cuatro grupos de cuatro subportadoras correspondientes a cada una de las celdas adyacentes sean ortogonales entre sí, esto es, no ocupen las mismas subportadoras. Además, en el caso de utilizar multiplexado espacial, el canal PCFICH se transmite por la misma antena que lo hace el canal PBCH.

En la región de control del canal PDCCH también se ubican los canales PHICH, dedicados a transmitir los reconocimientos (ACK/NACK) asociados a la transmisión de paquetes en el enlace ascendente. Utilizando secuencias ortogonales de Walsh (véase anexo 5.1) se pueden multiplexar múltiples canales PHICH sobre el mismo conjunto de recursos. En la Figura 5-9 se muestra una posible ubicación del canal PHICH. El número de símbolos utilizados para transmitir los grupos PHICH puede ser típicamente 1 ó 3. El canal PBCH indica el número de grupos PHICH configurados en la celda y el número de símbolos utilizados.

Tal como muestra la Figura 5-9, el canal PDCCH también contiene los denominados *mini Control Channel Elements* (mini-CCEs), que como se ha indicado en el apartado correspondiente a la descripción del canal PDCCH transmiten la información de cuatro parámetros (denominados *Resource Element Groups* o REGs) relacionados con la capa física.

El resto de los recursos en la rejilla frecuencia-tiempo de la subtrama no utilizados para transmitir el canal PDCCH se destinan en general a ubicar distintos canales PDSCH a partir de los mecanismos de asignación de recursos radio descritos en el capítulo 6 de este libro.

Finalmente, obsérvese en la Figura 5-9 que en determinadas posiciones de la rejilla frecuencia-tiempo de la subtrama se ubican las denominadas señales de referencia. Obviamente en estas posiciones no se puede transmitir ningún símbolo o bit asociado a los canales descritos anteriormente.

²² En el caso de transmitir señales MBMS en un sistema iso-frecuencia no se transmite el canal PCFICH ya que en este tipo de sistemas no disponen de canales físicos de control.

²³ El identificador de celda se obtiene a partir de las señales de sincronización primarias y secundarias.

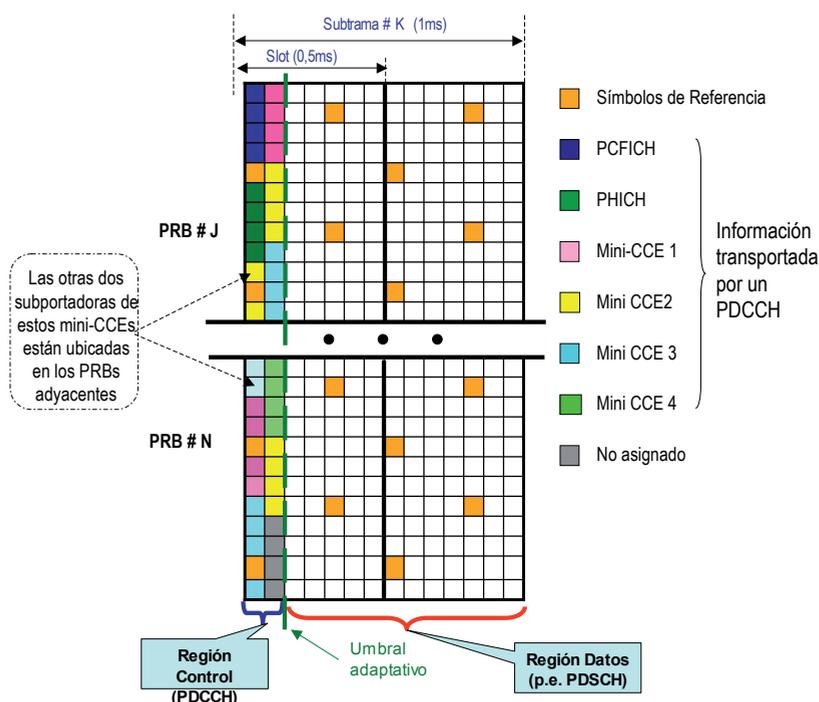


Figura 5-9 Detalle de la ubicación del contenido del canal PDCCH, considerando que ocupa sólo dos símbolos OFDMA

En el caso que se utilicen técnicas de multiplexado espacial (MIMO), los canales físicos comunes de control (PBCH, PDCCH) y las señales de sincronización se transmiten utilizando siempre el *port* (antena) numerado como cero, mientras que los canales físicos de tráfico PDSCH asignados a los diferentes usuarios pueden transmitirse utilizando el *port* cero o cualquiera de los otros que conforman la estructura MIMO. Recordar que en caso de utilizar estructuras MIMO las señales primarias y secundarias de referencia son específicas para cada *port*. Véase sección 5.5.3.1 de este capítulo.

Finalmente, hay que indicar que el sistema LTE está diseñado para tener un alto grado de compatibilidad entre los procedimientos a nivel de capa física aplicados en ambos modos FDD y TDD. O dicho de otra forma, la mayor parte de la descripción sobre el mapeo de los diferentes canales físicos en los recursos frecuencia-tiempo indicada anteriormente para el sistema FDD, también es válida para el sistema TDD. De hecho, y dejando aparte la distinta naturaleza del método de duplexado utilizado, la mayor diferencia reside en la ubicación de las señales de sincronización. En efecto ahora, cuando se opera en modo TDD, la señal primaria de sincronización aparece en el tercer símbolo de las subtramas especiales, mientras que la señal secundaria de sincronización aparece tres símbolos antes que la primaria. Además, hay

que subrayar que, tal como ocurre cuando se opera en modo FDD, el canal de radiodifusión (PBCH) en modo TDD también se sitúa en la primer sub-trama de la estructura de trama.

5.5.4 Enlace Ascendente

De modo análogo a lo realizado en el apartado 5.5.3, en este apartado se revisan las principales características de los procedimientos y canales físicos utilizados para transportar la información de usuario hacia el eNB, así como de los mecanismos de señalización y control que posibilitan una transmisión eficiente y controlada de la información generada por los usuarios. El apartado también incluye la descripción de las señales físicas que permiten al eNB sincronizar y demodular coherentemente las señales SC-FDMA transmitidas por los diferentes terminales móviles bajo el área de cobertura de dicho eNB. Finalmente, el apartado concluye indicando cómo los diferentes canales definidos están ubicados en los recursos frecuencia-tiempo disponibles.

5.5.4.1 Canales Físicos en el enlace ascendente:

En este apartado se describen brevemente las características más relevantes de los canales físicos especificados para el enlace ascendente en el sistema LTE. Como en el caso del enlace descendente se distingue entre canales físicos de tráfico compartidos y canales físicos de control, utilizados para transportar tanto señales de control generadas en la capa física del sistema como los mensajes de control generados en las capas superiores del sistema.

Physical Uplink Shared Channel (PUSCH)

Es el canal utilizado para enviar la información de usuario. Se transmite utilizando señales SC-FDMA²⁴ durante una o varias subtramas. El número de subportadoras utilizado lo determina el gestor de recursos radio (*scheduler*) del enlace ascendente. Esta asignación la hace el eNB y se transmite al terminal móvil mediante el canal PDCCH (véase Tabla 5-7).

Los esquemas de modulación considerados son: QPSK, 16QAM y 64 QAM (opcional), mientras que para la codificación de canal utiliza turbo códigos de tasa 1/3. Adicionalmente utiliza un mecanismo de retransmisión híbrido (H-ARQ) a partir de un código de redundancia cíclico (CRC) de 24 bits.

Cuando los informes sobre la calidad del canal no son suficientemente precisos, por ejemplo debido a un alto valor de la frecuencia Doppler del canal, el sistema puede hacer uso del amplio ancho de banda disponible y aplicar técnicas de diversidad en frecuencia, en concreto técnicas de salto de frecuencia (*Frequency Hopping*). Tal como se muestra en la Figura 5-10, se pueden aplicar dos tipos de técnicas de salto de frecuencia: entre subtramas o dentro de la misma subtrama. En el caso de saltos de frecuencia en la misma subtrama, el cambio de

²⁴ Véase capítulo 4 para detalles sobre la señal SC-FDMA

sub-portadoras asociadas al canal PUSCH se produce en la transición de una ranura temporal (TS) a otra, tal como muestra la Figura 5-10. Por su parte, el mecanismo de salto de frecuencia entre subtramas proporciona diversidad en frecuencia entre las retransmisiones HARQ de un mismo paquete de información, ya que la asignación de portadoras cambia en cada asignación de subtrama. Obsérvese que en este caso la retransmisión se efectúa 8 subtramas más tarde con respecto a la transmisión inicial tal como especifica el mecanismo de retransmisión HARQ síncrono, propio del enlace ascendente (véase anexo 5.3 de este capítulo).

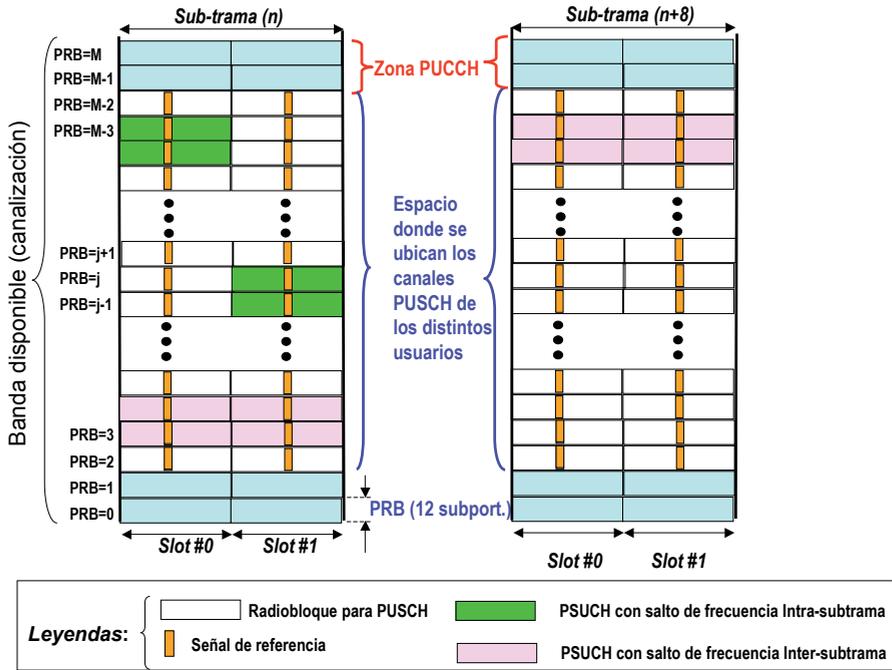


Figura 5-10 Procedimientos de diversidad en frecuencia en el enlace ascendente

En ambos casos el patrón de salto de frecuencia a aplicar depende de la identidad de la celda (*Cell-Id*) y se indica al terminal móvil mediante el canal de señalización PDCCH del enlace descendente.

Finalmente, es importante subrayar que el canal PUSCH no es un canal dedicado, sólo se asigna al usuario cuando éste tiene algo que transmitir en el enlace ascendente (UL), de acuerdo con un cierto algoritmo de *scheduling*, y en base a las peticiones que dicho usuario haga a través del canal de control PUCCH descrito a continuación.

CANALES DE CONTROL

Por lo que respecta a los canales físicos de control del enlace ascendente, se clasifican en:

Physical Uplink Control Channel (PUCCH)

Physical Random Access channel (PRACH)

Physical Uplink Control Channel (PUCCH)

Contiene información de control del enlace ascendente. En particular puede transmitir las siguientes informaciones de control:

- Peticiones de asignación de recursos (*Scheduling Request*).
- Reconocimientos (ACK/NACK) para el mecanismo de retransmisión híbrido (HARQ) de paquetes en el enlace descendente.
- Información de la calidad del canal (*Channel Quality Indicator-CQI*) necesaria para optimizar los procedimientos de asignación de recursos radio en el enlace descendente.

De forma general cabe indicar que el canal PUCCH se ubica en los extremos de la banda asignada al sistema, tal como se muestra en la Figura 5-11, en donde se muestra un ejemplo con cuatro canales PUCCH (numerados de 0 a 3), cada uno de ellos ocupando un radiobloque físico (PRB). Esta ubicación impone pocas limitaciones en la asignación de recursos en el enlace ascendente y permite que el canal de datos de usuario (PUSCH) pueda ocupar toda la parte central de la banda disponible. Además, con el fin de dotarle de un mecanismo de diversidad en frecuencia, que le proporcione robustez frente a desvanecimientos, los canales PUCCH aparecen duplicados (es decir se envía la misma información) y su posición intercambiada entre la primera y segunda ranuras temporales (TS) de la subtrama, tal como se indica en la Figura 5-11. La ganancia aportada por este mecanismo de diversidad en frecuencia es del orden de 2 dB

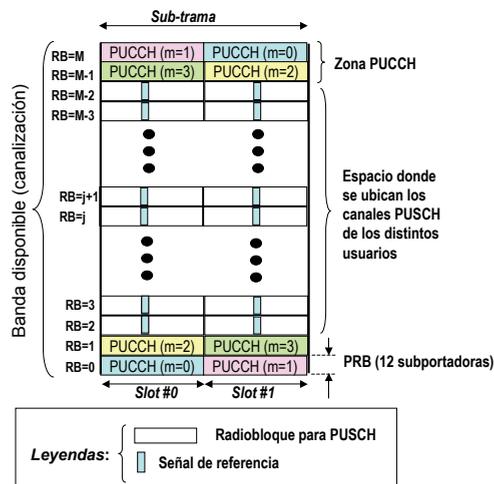


Figura 5-11 Ejemplo de ubicación de los canales PUCCH dentro de la banda de canalización del sistema

El número de radiobloques utilizados para transmitir el canal PUCCH se indica en el canal de radiodifusión (PBCH) y depende de la canalización considerada. En la Tabla 5-10 se indican valores típicos para las diferentes canalizaciones estandarizadas

Tabla 5-10. - Número de PRBs típico asignados al canal PUCCH

Canalización (MHz)	Número de PRBs asignados	Número de canales PUCCH
1.4	2	1
3	4	2
5	8	4
10	16	8
20	32	16

Obsérvese que debido al mecanismo de diversidad en frecuencia el número de PRBs utilizado es el doble que el número de canales PUCCH transmitidos

Es conveniente subrayar que en un canal PUCCH es posible transmitir información de control de múltiples usuarios simultáneamente. Para ello se utiliza una técnica de multiplexación ortogonal por división en el código (CDM), bien en frecuencia bien en tiempo o en los dos dominios, para mezclar las diferentes informaciones de señalización transmitidas.

El canal PUCCH utiliza típicamente modulación BPSK y QPSK y se transmite utilizando diferentes tipos de formato, tal como se indica en la Tabla 5-11.

Tabla 5-11 Configuración del canal PUCCH

Formato del canal PUCCH	Esquema de modulación	Número de bits por subtrama	Tipo de información Transmitida
1	ON/OFF	---	Petición de asignación de recursos (scheduling)
1a	BPSK	1	1 bit ACK/NACK para HARQ
1b	QPSK	2	2 bit ACK/NACK para HARQ
2	QPSK	20	CQI (20 bits codificados)
2 (mod.)	QPSK	20	CQI más 1 ó 2 bits ACK/NACK para prefijo cíclico largo
2a	BPSK+QPSK	21	CQI (20 bits) más 1 bit ACK/NACK
2b	BPSK+QPSK	22	CQI (20 bits) más 2 bit ACK/NACK

En el formato tipo 1, los bits de información de control, previamente aleatorizados, se multiplican bit a bit por dos secuencias, una primera de Zadoff-Chu de longitud 12 desplazada cíclicamente (lo que proporciona ortogonalidad en el dominio de la frecuencia entre terminales móviles situados en la misma celda), seguida de una secuencia ortogonal de Walsh-Hadamard²⁵, (lo que proporciona ortogonalidad en el dominio del tiempo). Para el formato tipo 2 sólo se utiliza la secuencia Zadoff-Chu que caracteriza al eNB.

Transmisión de los ACK/NACK del mecanismo de retransmisión híbrido (formatos 1a/1b)

Cuando se utiliza el canal físico PUCCH sólo para transmitir información de reconocimientos (ACK/NACK) correspondientes al mecanismo de retransmisión híbrido del enlace descendente, el procedimiento de generación de la información transmitida por el canal PUCCH es el siguiente:

- En el caso de considerar que el reconocimiento está compuesto por un único bit, se utiliza modulación BPSK. Si el reconocimiento está compuesto por 2 bits se utiliza modulación QPSK.
- La señal modulada pasa por un codificador de Walsh que proporciona ortogonalidad en el dominio del tiempo y permite multiplexar diferentes usuarios en el este dominio. A su salida se dispone de "N" ramas en paralelo, tantas como símbolos tenga la secuencia de Walsh utilizada.
- En cada rama el símbolo modulado es multiplicado por una secuencia de Zadoff-Chu de longitud 12 desplazada cíclicamente (véase anexo 5.1, sección I-1). Posteriormente la señal resultante se inyecta a un bloque que efectúa la transformada discreta de Fourier (DFT) para generar doce subportadoras moduladas BPSK/QPSK. Cada uno de estos grupos de 12 subportadoras se transmiten utilizando un símbolo SC-FDMA de la subtrama.
- Los símbolos de referencia utilizados para la estimación de la respuesta impulsional del canal se ubican en medio de la ranura temporal (i.e., *slot*). En el caso de utilizar un prefijo cíclico normal, los símbolos numerados como #2, #3 y #4 llevan dichas señales, mientras que si se utiliza el prefijo cíclico extendido o largo sólo se ubican dos símbolos de referencia en las posiciones #2 y #3.

Véase en la Figura 5-12 el diagrama de bloques de la generación de los símbolos SC-FDMA de la señal PUCCH en caso de utilizar los formatos 1a/1b, y en la Figura 5-13 la ubicación de las señales de referencia y ACK/NACK para un canal PUCCH con los mencionados formatos.

²⁵ En el anexo 5.1 se describen las propiedades de las secuencias de Walsh-Hadamard.

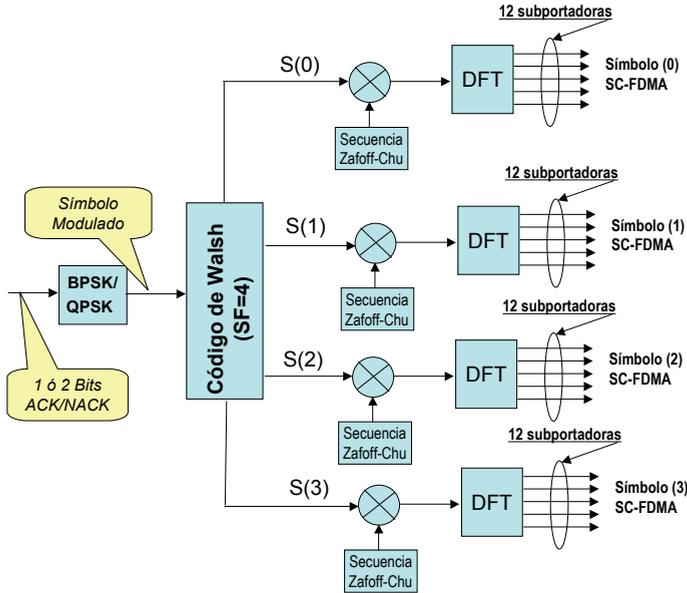


Figura 5-12.- Generación de símbolos SC-FDMA para transmisión del canal PUCCH en formatos 1a/1b. Ejemplo con secuencia de Walsh de longitud 4

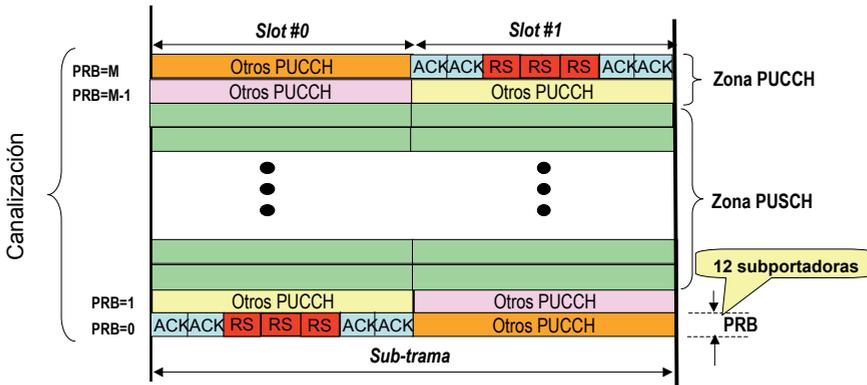


Figura 5-13.- Ubicación de las señales de referencia y ACK/NACK en el canal PUCCH en formatos 1a/1b

Aprovechando las propiedades de correlación de las secuencias Zadoff-Chu²⁶ y la ortogonalidad proporcionada en el dominio temporal por las secuencias de Walsh, se puede multiplexar en la misma subtrama información de reconocimiento de hasta 18 usuarios en el mismo PUCCH. En efecto, en cada ranura temporal (TS) se pueden considerar entre 4 y 6 se-

²⁶ Recordar (véase anexo 5.1) que la transformada de Fourier (IFFT) de una secuencia de Zadoff-Chu generada en el dominio temporal es otra secuencia de Zadoff-Chu, con idénticas propiedades de correlación en el dominio frecuencial

cuencias de Zadoff –Chu desplazadas cíclicamente, lo que proporciona ortogonalidad en el dominio de la frecuencia, mientras se pueden utilizar códigos de Walsh-Hadamard con factor de ensanchamiento 2 y 4, para proporcionar ortogonalidad en el dominio temporal.

Transmisión de los indicadores de calidad del canal CQI (formato 2)

Cuando se utiliza el canal físico PUCCH para transmitir sólo información sobre la calidad del canal radio, el indicador CQI consta de entre 5 y 10 bits, dependiendo del tamaño del informe. Típicamente se transmite utilizando una subtrama de 1 ms, aunque para informes largos puede utilizar múltiples subtramas. El procedimiento de generación de la información del canal PUCCH es el siguiente:

- Los 10 bits del informe CQI se codifican con tasa $\frac{1}{2}$ mediante un código de Reed-Muller para generar 20 bits codificados, que son aleatorizados mediante una secuencia de Gold.
- Los 20 bits aleatorizados son modulados QPSK para dar lugar a 10 símbolos, que, una vez multiplicados cada uno de ellos por una secuencia de Zadoff-Chu de longitud 12 desplazada cíclicamente dan lugar, a partir de una transformada discreta de Fourier (DFT), y para cada símbolo QPSK original, a diez grupos de doce subportadoras moduladas QPSK. Cada uno de estos grupos de 12 subportadoras se transmiten utilizando un símbolo SC-FDMA de la subtrama.
- En el caso de utilizar un prefijo cíclico normal, las señales de referencia utilizadas en la estimación de la respuesta impulsional del canal se ubican en los símbolos numerados como #1 y #5 de cada una de las ranuras temporales (TS) de la subtrama. En caso de utilizar prefijo cíclico extendido o largo, las señales de referencia sólo se ubica el símbolo numerado como #4.

Véase en la Figura 5-14 el diagrama de bloques correspondiente a la generación de los símbolos SC-FDMA de la señal PUCCH en caso de utilizar el formato 2, y en la Figura 5-15 para la ubicación de las señales de referencia y CQI en el canal PUCCH con formato 2.

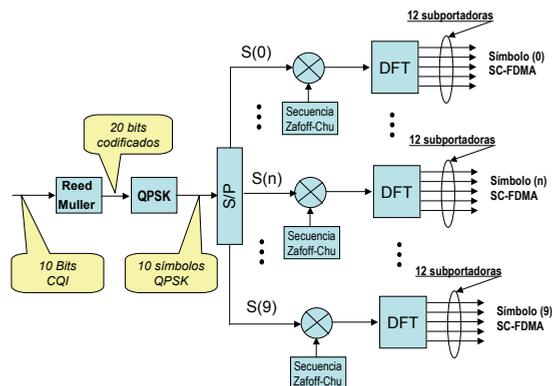


Figura 5-14.- Generación de símbolos SC-FDMA para transmisión del canal PUCCH en formato 2

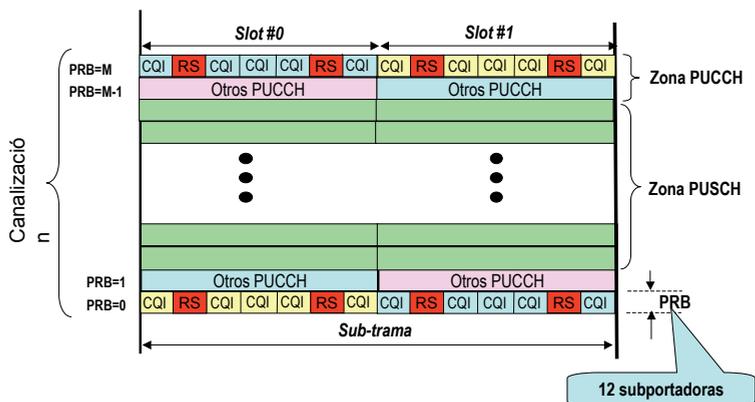


Figura 5-15 Ubicación de las señales de referencia y CQI en el canal PUCCH en formato 2

Nuevamente, aprovechando las propiedades de correlación de las secuencias Zadoff-Chu, se pueden multiplexar en la misma subtrama los informes de hasta 12 usuarios en el mismo PUCCH.

Multiplexado de los indicadores de calidad del canal CQI y ACK/NACK (formatos 2a/2b)

La estructura básica de transmisión para estos formatos es la misma que la utilizada para el formato 2. La única diferencia es que ahora también hay que ubicar a la señal ACK/NACK que puede ser de 1 ó 2 bits. Para ello primero se modula dicha señal en BPSK (caso de 1 bit) o QPSK (caso de 2 bits). El símbolo resultante (d_{HARQ}) se ubicará en el símbolo SC-FDMA numerado como #5 de cada ranura temporal (TS), substituyendo a la segunda señal de referencia. Para ello dicho símbolo, d_{HARQ} , se multiplicará por la correspondiente secuencia de Zadoff-Chu de longitud 12 desplazada cíclicamente, quién una vez efectuada la transformada discreta de Fourier (DFT) dará lugar a las doce subportadoras moduladas BPSK/QPSK que se transmitirán en el mencionado símbolo SC-FDMA.

Physical Random Access Channel (PRACH)

El canal físico de acceso aleatorio consta de un prefijo cíclico y un preámbulo. El preámbulo utiliza secuencias de Zadoff-Chu por sus buenas propiedades de correlación cruzada. La señal así resultante es modulada SC-FDMA para ser transmitida en los denominados bloques de acceso, definidos como un conjunto de 72 sub-portadoras contiguas ubicadas, por el eNB para dicho fin, dentro de la estructura frecuencia tiempo de enlace ascendente. Véase Figura 5-21.

Las secuencias preámbulo utilizadas parten de una secuencia de Zadoff-Chu raíz (o secuencia inicial) de longitud 839, y se generan por desplazamiento cíclico de dicha

secuencia raíz. Existen un total de 838 posibles secuencias con óptima correlación cruzada²⁷. No obstante en la práctica sólo se consideran como mucho 64, generadas con una separación cíclica suficientemente grande. Ello es debido a que en recepción se van a utilizar procedimientos de correlación para detectar ésta secuencia preámbulo. En efecto, haciendo uso del hecho que la transformada Discreta de Fourier (DFT) de una secuencia de Zadoff-Chu es otra secuencia de Zadoff-Chu ponderada y desplazada cíclicamente (véase anexo 5.1), y del hecho que los diferentes preámbulos utilizados en el canal PRACH proceden de una misma secuencia raíz desplazada cíclicamente, en recepción la señal resultante a la salida del correlador (trabajando ya en el dominio temporal) tendrá una imagen tal como la mostrada en la Figura 5-16, en donde se aprecia la señal a la salida del correlador asociada al preámbulo hipotéticamente transmitido y los picos de correlación de otros dos posibles preámbulo (generados por un desplazamiento N_{CS} de la misma secuencia raíz). Obsérvese además que utilizando las excelentes propiedades de correlación de la secuencia de Zadoff-Chu también es posible estimar la respuesta impulsional del canal, de hecho es la señal que se aprecia a la salida del correlador resultante de la detección del preámbulo PRACH. Por lo tanto, y dado que existe incertidumbre temporal asociada al retardo de propagación (inicialmente desconocido ya que se desconoce la distancia entre el terminal móvil y el eNB), el desplazamiento N_{CS} entre secuencias preámbulo debe ser mayor que el intervalo temporal asociado al máximo retardo de propagación posible más la dispersión temporal del canal, tal como se muestra en la Figura 5-16.

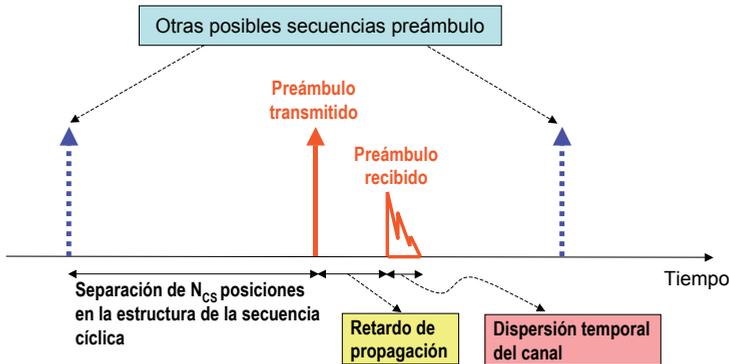


Figura 5-16. Efecto en el eje temporal de la separación cíclica entre posibles secuencias preámbulo

En el caso de operar con la estructura de trama tipo 1 (véase Figura 5-3) se definen cuatro posibles formatos que se diferencian en la duración del prefijo cíclico (T_{CP}) y del preámbulo (T_{preamb}), mientras que en el caso de operar con una estructura de trama tipo 2 (véase Figura 5-4) sólo hay tres formatos posibles. Para la estructura de trama tipo 1 (SF1) hay 64 secuencias preámbulo disponibles por cada celda, mientras que para la estructura de trama tipo 2 sólo hay 16. En la Tabla 5-12 se resumen los posibles formatos del canal PRACH.

²⁷ Véase anexo 5.1 sobre propiedades de las secuencias de Zadoff-Chu.

Tabla 5-12 Estructuras posibles del canal PRACH

Estructura Trama	Formato	T_{CP}	T_{preamb}
Tipo 1	0	0,1ms	0,8ms
	1	0,68ms	0,8ms
	2	0,2ms	1,6ms
	3	0,68ms	1,6ms
Tipo 2	0	0ms	0,133ms
	1	0ms	0,533ms
	2	N/A	NA

El proceso de acceso aleatorio implica tanto a la capa física como a las capas superiores. Desde el punto de vista de la capa física el procedimiento de acceso a la red se inicia con la transmisión por parte del terminal móvil de un preámbulo de acceso aleatorio. Si el eNB detecta correctamente este preámbulo, envía mediante el canal PDSCH una respuesta con, entre otros, un identificador del recurso tiempo-frecuencia en donde el preámbulo RACH se detectó (con ello identificamos al móvil al que va destinada la respuesta), el valor inicial del *Time Advance* (véase sección 5.5.4.2) y una asignación inicial de recursos frecuencia-tiempo para poder continuar el procedimiento de acceso (véase sección 5.8.2). En caso contrario, al cabo de un cierto tiempo aleatorio el móvil repite la operación aumentando la potencia de transmisión (mecanismo de *Power Ramping*) y así sucesivamente hasta que el eNB detecta el preámbulo.

Por lo que respecta a las capas superiores del terminal móvil, éstas eligen, entre las subportadoras disponibles, el conjunto (denominado bloque de acceso) que se va a utilizar para enviar el preámbulo de acceso. Además las capas superiores envían a la capa física la siguiente información:

- Bloques de acceso disponibles.
- Formato del preámbulo (qué secuencia utilizar).
- Potencia inicial de transmisión.
- Número máximo de reintentos.
- Perfil de aumento/disminución de la potencia (*power ramping*).

5.5.4.2 Mecanismo de Avance Temporal (Time Advance)

Para garantizar que las transmisiones efectuadas en el enlace ascendente por los distintos usuarios dentro de una trama no se interfieran entre si, hay que garantizar que el inicio de todas ellas llega al eNB en un intervalo temporal, medido desde el inicio teórico de las distintas

subtramas, inferior a la duración del prefijo cíclico de un símbolo SC-FDMA. Ello requiere de un estricto mecanismo de sincronización que permita compensar los diferentes tiempos de propagación asociados a usuarios más o menos alejados del eNB. A este mecanismo se le denomina mecanismo de avance temporal o *Time Advance*. La filosofía del mecanismo es simple. Se trata de que el terminal móvil inicie su transmisión en el enlace ascendente con un cierto avance temporal (cuyo valor es proporcionado por el eNB) con respecto al inicio teórico de la misma, de modo que se compense el retardo asociado a la propagación, tal como se detalla a continuación.

En la Figura 5-17 se muestra el principio de funcionamiento básico de dicho mecanismo.

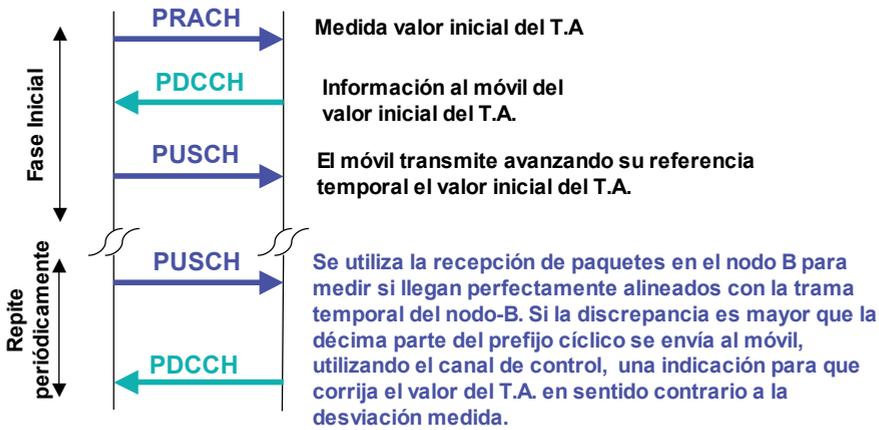


Figura 5-17 Mecanismo de Avance temporal

Puesto que el terminal móvil puede cambiar de ubicación, la distancia entre su posición y la del eNB puede variar. Por lo tanto hay que ajustar permanentemente el valor del Avance Temporal para corregir y compensar los cambios en el valor del tiempo de propagación. Por consiguiente, en el mecanismo de avance temporal hay que diferenciar dos situaciones:

- Adquisición del valor inicial del Avance Temporal. Para ello se utiliza el preámbulo de acceso aleatorio (PRACH). Puesto que este es la primera señal que transmite el terminal móvil, en el momento de su transmisión no se tiene conocimiento del valor apropiado del valor del Avance Temporal a utilizar. Por consiguiente, este preámbulo inicia su transmisión según el patrón temporal del terminal móvil, y en consecuencia llegará al eNB con un retardo igual al $2T_{\text{pro}}^{28}$ respecto patrón temporal de dicho nodo, tal como se indica en la Figura 5-18.

²⁸ T_{pro} indica el retardo debido a la propagación, por consiguiente el patrón temporal del móvil está retardado T_{pro} con respecto al existente en el eNB. Puesto que el preámbulo de acceso se transmite utilizando el patrón temporal del móvil, éste llegará al nodo eNB con un retardo igual a $2T_{\text{pro}}$. Por lo tanto, el valor $2T_{\text{pro}}$ mide la diferencia entre el instante teórico de la subtrama y el momento de la llegada del preámbulo de acceso aleatorio, o dicho de otra forma, si se desea que el resto de las transmisiones del móvil lleguen al nodo eNB correctamente alineadas temporalmente, éstas deberán iniciarse $2T_{\text{pro}}$ segundos antes del inicio temporal de la subtrama indicada según el patrón del terminal del móvil. Véase Figura 5-18 para detalles.

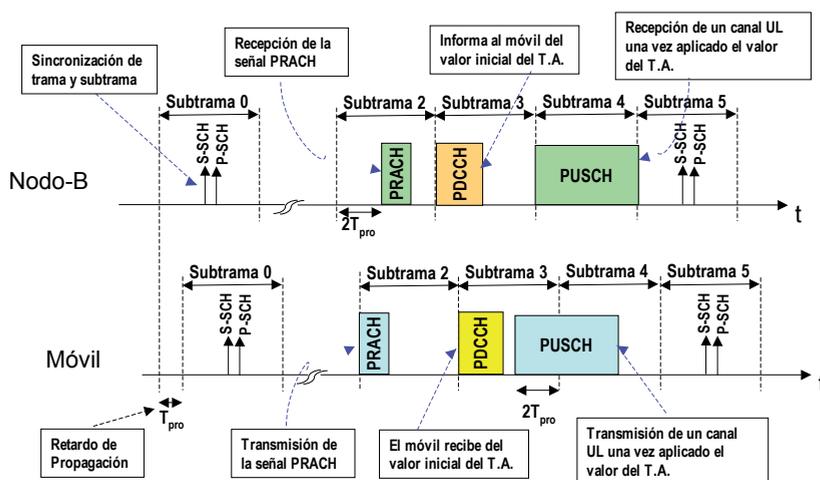


Figura 5-18 Medida inicial del valor del Avance temporal

Es decir, el eNB podrá medir el valor inicial del tiempo de propagación y aprovechando la respuesta a la petición de acceso (véase sección 5.8.2) envía al móvil el valor del avance temporal a aplicar, es decir el valor $2T_{prop}$, con una resolución de $0,52 \mu s$ (una décima parte de la duración del prefijo cíclico normal).

- Avance temporal en régimen permanente. Una vez establecido el valor del avance temporal inicial, en las siguientes transmisiones del enlace ascendente el eNB va midiendo la desviación temporal con que le llegan los paquetes. Si esta desviación temporal es mayor de $\pm 0,52 \mu s$. envía al móvil, utilizando el canal PDCCH, un comando de corrección indicándole que aumente o disminuya, según convenga, el valor actual del Avance Temporal en $0,52 \mu s$ para garantizar que los paquetes lleguen al eNB correctamente alineados con el inicio del prefijo cíclico.

5.5.4.3 Señales de referencia en UL

De modo análogo a lo que ocurre en el enlace descendente, también se definen dos tipos de señales de referencia para el enlace ascendente:

- Señales utilizadas en la estimación de la respuesta al impulso del canal, denominadas *DeModulation Reference Signal* o DM-RS, necesarias para permitir la demodulación coherente en el enlace ascendente²⁹.
- Señales de referencia destinadas al sondeo de la calidad del canal de transmisión, denominadas SRS (i.e., *Sounding Reference Signals*).

²⁹ Los mecanismos de estimación de la respuesta al impulso del canal no están estandarizados ya que se deja a cada fabricante de equipos que pueda implementar los suyos propios. Véase referencia [16] para un resumen de las posibles técnicas

Las señales de referencia tipo DM-RS están asociadas a la transmisión de datos de usuario o información de control, es decir a la transmisión de paquetes PUSCH o PUCCH. Cuando se considera la transmisión de paquetes PUSCH y en el caso que el sistema utilice prefijo cíclico normal entonces la señal DM-RS siempre aparece en el cuarto símbolo SC-FDMA de todas las subtramas transmitidas por el usuario y ocupa el mismo número de subportadoras que el bloque de datos de usuario. En caso que el sistema utilice un prefijo cíclico largo, entonces la señal DM-RS aparece en el tercer símbolo SC-FDMA de cada subtrama. En el caso de considerar paquetes PUCCH la posición de las señales de referencia no es fija y depende del formato utilizado³⁰.

La señal DM-RS transporta un símbolo SC-FDMA compuesto por $N \times 12$ subportadoras (donde N indica el número de PRBs asignados al canal PUSCH o PUCCH) cuyo contenido se corresponde a una secuencia Zadoff-Chu (en el dominio de la frecuencia) de idéntica longitud, secuencia que es utilizada para estimar la respuesta impulsional. Cada celda utiliza una secuencia de referencia raíz diferente, mientras que para diferenciar usuarios de una misma celda las secuencias de referencia utilizadas se generan a partir de desplazamientos cíclicos de la secuencia raíz de la celda.

A diferencia de las señales de referencia DM-RS, las señales SRS no están asociadas a ninguna transmisión de señales de datos y/o de control en el enlace ascendente. Se utilizan fundamentalmente para que la red pueda conocer el comportamiento del canal en el dominio de la frecuencia y facilitar de esta forma la decisión de los algoritmos de gestión de recursos radio (*scheduling*). No obstante también pueden ser utilizadas para mejorar los mecanismos de control de potencia o incluso facilitar los mecanismos de configuración inicial del terminal móvil. Estas señales son transmitidas por los terminales móviles en aquellas subtramas que la red les indica mediante mensajes de radiodifusión. La asignación puede ser para una subtrama individual o de forma periódica. Los periodos establecidos en el estándar LTE son 2, 5, 10, 20, 40, 80, 160 ó 320 ms.

Las señales SRS siempre se transmiten en el último símbolo de las subtramas seleccionadas. Puesto que múltiples terminales móviles pueden necesitar enviar señales SRS hacia el eNB, se utiliza una técnica de asignación de subportadoras basada en el denominado *Interleaved FDMA* que asigna de forma alternada subportadoras a los diferentes móviles, tal como se muestra en la Figura 5-19.

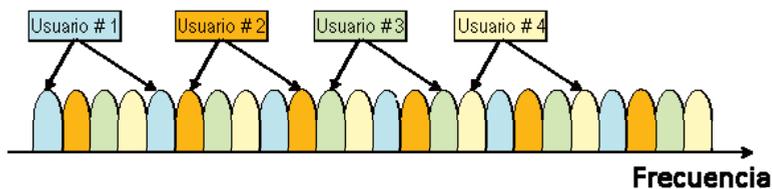


Figura 5-19 Asignación de subportadoras para la señal SRS. Ejemplo para 4 móviles

³⁰ Véase sección 5.5.4.1 apartado correspondiente al canal PUCCH (Figura 5-13, Figura 5-15) para detalles.

Finalmente indicar que las señales SRS también utilizan secuencias de Zadoff-Chu, del mismo tipo que las utilizadas para la señal de referencia DM-RS.

5.5.4.4 Mapeo de los canales Físicos

Antes de proceder al mapeo de los diferentes canales físicos utilizados en el enlace ascendente, y a modo de recordatorio, en la Tabla 5-13 se describen dichos canales.

Tabla 5-13 Resumen de los canales físicos del enlace ascendente

Nombre Canal	Descripción
<i>Physical Uplink Shared Channel (PUSCH)</i>	Transmite información de usuario y de control dedicada procedente de capas superiores a la capa física.
<i>Physical Uplink Control Channel (PUCCH)</i>	Contiene información de control del enlace ascendente. En particular transmite: Peticiones de asignación de recursos (<i>Scheduling Request</i>); Reconocimientos (ACK/NACK) correspondientes al enlace descendente, información de la calidad del canal (<i>Channel Quality Indicator-CQI</i>).
<i>Physical Random Access Channel (PRACH)</i>	Canal físico de acceso aleatorio. Envía un preámbulo para iniciar, por parte del móvil, el procedimiento de conexión al sistema..

En la Figura 5-20 se indica cómo se ubican en la estructura frecuencia-tiempo los diferentes canales físicos del enlace ascendente para el caso de FDD. Como se observa en la figura en los radiobloques extremos de la banda asignada se ubica el canal físico de control (PUCCH). Ello permite explotar el mecanismo de diversidad en frecuencia ya que, tal como se indicó en la Figura 5-11, los campos de cada información PUCCH de cada subtrama se cruzan a nivel de *slot*. Recuérdese que una subtrama está compuesta por 2 slots.

Los símbolos de referencia utilizados para la estimación de la respuesta impulsional del canal (indicados como DM-RS en la figura), y que permiten la demodulación coherente en el enlace ascendente del canal PUSCH, se transmiten en los símbolos s_3 de cada uno de los dos *slots* que componen la subtrama. Esta señal (símbolos) de referencia ocupa el mismo número de subportadoras que el canal PUSCH. En las subportadoras no asignadas a ningún usuario (recursos en blanco en la Figura 5-20) tampoco aparecen los correspondientes símbolos de referencia (recursos en gris en la Figura 5-20 y están indicados como RS-No utilizada).

En el último símbolo de cada una de las subtramas asignadas a tal fin y en las subportadoras no ocupadas por el canal PUCCH se ubican las señales de referencia (SRS) destinadas al sondeo de la calidad del canal de transmisión. En la figura se aprecia como esta señal ocupa todas las subportadoras asignadas a la canalización del sistema (excepto las correspondientes al canal PUCCH).

El resto de los recursos frecuencia-temporales de la subtrama pueden utilizarse para transmitir el canal de datos de PUSCH.

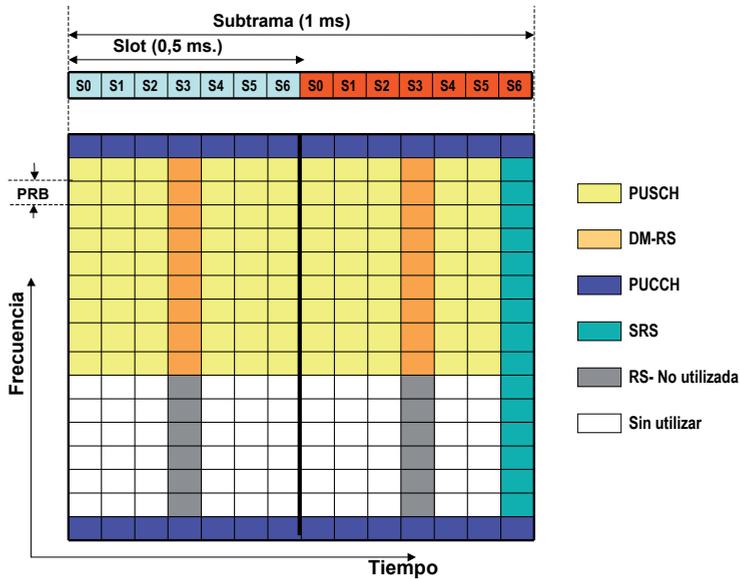


Figura 5-20 Ubicación en la estructura frecuencia-tiempo de los canales de control, canales de datos y señales de referencia para el enlace ascendente

En la Figura 5-20 no se indica la posición del canal de acceso aleatorio PRACH ya que son las capas superiores quienes fijan dicha posición en estructura frecuencia-tiempo. En todo caso, es importante mencionar que dicha localización es semiestática y se repite periódicamente dentro de la región dedicada a los canales tipo PUSCH, tal como se muestra en Figura 5-21.

Finalmente, y tal como ocurría en el enlace descendente, también en el enlace ascendente la mayoría de canales físicos se mapean de forma similar para ambas estructuras de duplexación FDD y TDD. La mayor diferencia reside en que en el caso de TDD es posible ubicar un preámbulo de acceso aleatorio corto en el intervalo U_pPTS de la subtrama especial. En efecto en modo TDD el estándar define un preámbulo de acceso aleatorio corto (*Short-RACH* o *S-RACH*) compuesto por una secuencia de Zadoff-Chu de longitud 139 de duración $133\mu s$. Este preámbulo ocupa los dos símbolos del U_pPTS de la subtrama especial. Es importante mencionar que además del *S-RACH* el estándar también permite en modo TDD la transmisión de preámbulos PRACH en subtramas convencionales asignadas al enlace ascendente. Como en el caso FDD el número y la ubicación de estas señales no es fijo y está definido por las capas superiores.

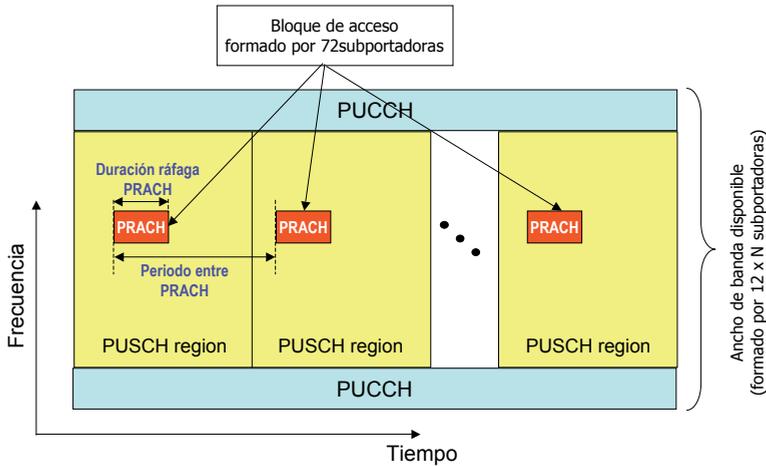


Figura 5-21 Ubicación del canal PRACH

Además, en el caso de duplexado TDD las señales de referencia destinadas al sondeo del canal (SRS) no sólo se transmiten utilizando el último símbolo de las subtramas a ello destinadas al enlace ascendente sino que también pueden transmitirse en uno o ambos símbolos asociados al campo UpPTS de la subtrama especial. Puesto que los símbolos UpPTS no pueden utilizarse para transmitir datos de usuario, y en tanto que siempre están presentes en una trama TDD, la transmisión de las señales SRS en dichos símbolos es prioritaria, salvo que se utilicen para transmitir preámbulos de acceso aleatorio cortos.

5.6 Mapeo entre canales Físicos, de Transporte y Lógicos

Enlace descendente

En la Figura 5-22 se muestra la relación entre los canales lógicos, de transporte y físicos en el caso del enlace descendente, [10]. En el caso del canal lógico de radiodifusión (BCCH), y dependiendo del tipo de información a transmitir, dicho canal puede transmitirse utilizando bien el canal de transporte de radiodifusión (BCH), bien el canal compartido (DL-SCH). En efecto, tal como se verá posteriormente en la sección 5.8.1, dedicada a describir los mecanismos de sincronización y adquisición inicial de información, el canal de de transporte de radiodifusión (BCH) sólo lleva la información del sistema mínima necesaria para que el terminal móvil pueda acceder al mismo (por ejemplo la canalización disponible en el sistema, el número de trama, o información de configuración de los canales físicos). El resto de la información sobre el sistema llega al terminal móvil mediante los denominados *System Information Block* (SIB)³¹ que se transmiten utilizando canal de transporte compartido (DL-SCH).

También es importante subrayar que dicho canal de transporte compartido (DL-SCH) se

³¹ Véase anexo 5.2 de este capítulo.

utiliza no sólo para transportar datos de usuario (DTCH) sino también señalización común (CCCH) y dedicada (DCCH). Ello es así ya que a diferencia de los sistemas de comunicaciones móviles previos, el sistema LTE no define canales físicos ni de transporte dedicados. En otras palabras, sólo cuando el usuario tiene datos o señalización a recibir o bien la red necesita enviar información de control común, se activa el canal compartido y por lo tanto no se desperdician recursos manteniendo activos canales dedicados que no siempre se utilizan.

Por lo que respecta a la transmisión de la información y/o señalización *multicast* (canales lógicos MTCH/MCCH), si bien en el caso de transmisión de este tipo de servicio en una única celda cabe utilizar también el canal de transporte compartido (DL-SCH), generalmente se utiliza un canal de transporte específico para este tipo de información (MCH) ya que habitualmente este tipo de servicio se implementa utilizando redes iso-frecuencia y por lo tanto disponer de un canal específico para dicho fin facilita el despliegue y gestión de este tipo de redes.

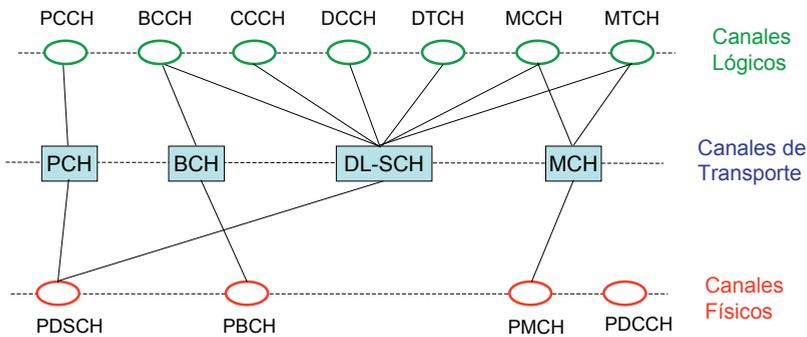


Figura 5-22 Relación entre los canales lógicos, de transporte y físicos del enlace descendente

En términos de canales físicos, es decir aquellos que implementan los mecanismos de transmisión a través del medio radioeléctrico, el sistema LTE es muy simple, ya que existe una relación directa entre el canal de transporte de radiodifusión (BCH) y su correspondiente canal físico (PBCH) o entre el canal de transporte *multicast* (MCH) y su correspondiente canal físico (PMCH). También existe una relación directa entre el canal de transporte compartido (DL-SCH) y su correspondiente canal físico (PDSCH) si bien este último también puede llevar el canal de transporte de aviso o *paging* (PCH). Puesto que el canal físico compartido (PDSCH) puede llevar muchos tipos de información y por lo tanto tener formatos diferentes, es necesario definir a nivel de capa física un canal de control (PDCCH) que lleve la señalización necesaria para procesar (modular, y demodular y ubicar en la estructura frecuencia-tiempo propia de la señal OFDMA) el mencionado canal PDSCH. Obviamente este canal PDCCH no tiene ninguna relación ni con los canales de transporte ni los canales lógicos, ya que sólo transmite señalización de capa física.

Enlace ascendente

En el caso del enlace ascendente la relación entre los canales lógicos, de transporte y

físicos, [10], es la indicada en la Figura 5-23. Para este enlace la interrelación es bastante más simple, por cuanto a nivel de canales lógicos sólo están definidos los canales de tráfico (DTCH), de control dedicado (DCCH) y de control común (CCCH) que, de forma análoga al enlace descendente, también se transportan utilizando un canal de transporte compartido, denominado ahora UL-SCH. Como en el caso del enlace descendente también ahora hay una relación unívoca entre dicho canal de transporte compartido y su correspondiente canal físico (PUSCH). Análogamente también se necesita definir un canal de control (PUCCH) a nivel de capa física que lleve la señalización necesaria para procesar (modular, demodular y ubicar en la estructura frecuencia-tiempo) la señal SC-FDMA que se utiliza en el mencionado canal PUSCH. Obviamente, tampoco este canal de control PUCCH no tiene ninguna relación ni con los canales de transporte ni los canales lógicos, ya que sólo lleva señalización de capa física.

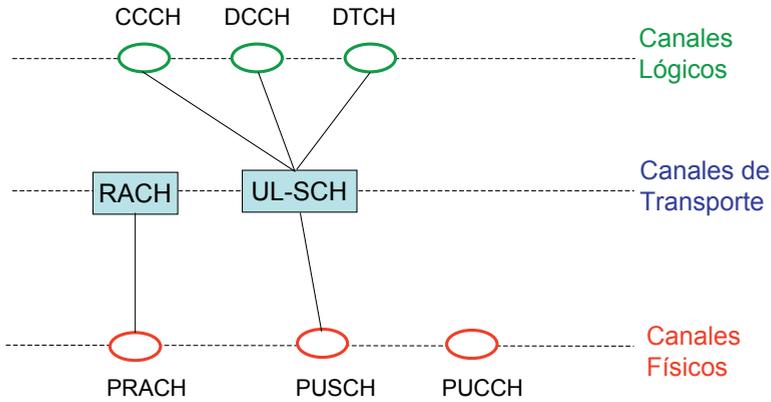


Figura 5-23 Relación entre los canales lógicos, de transporte y físicos del enlace ascendente

Por lo que respecta al canal de acceso aleatorio, obsérvese que en la Figura 5-23 sólo están definidos los canales físico (PRACH) y de transporte (RACH) que están relacionados de manera unívoca. No se precisa interacción a nivel de canal lógico ya que básicamente el canal de acceso es una secuencia preámbulo que sirve para informar al eNB de que un terminal móvil quiere conectarse al mismo. El resto de la señalización necesaria en el proceso de acceso aleatorio, por ejemplo la petición de conexión a nivel RRC (*RRC Connection Request*), actualización del área de localización (*Tracking area update*) o petición de asignación de recursos (*Scheduling Request*) puede ya transmitirse utilizando el canal compartido (UL-SCH) y por lo tanto, no es necesario establecer ningún otro canal lógico que realice la misma funcionalidad.

5.7 Categorías de terminales móviles en LTE

Se han definido, [10], cinco categorías diferentes de terminal móvil dependiendo de la capacidad MIMO, de la máxima velocidad de transmisión de pico y de la capacidad de procesado del terminal móvil. En la Tabla 5-14 se resumen las características más relevantes de las

diferentes categorías tanto para el caso del enlace descendente como ascendente.

Tabla 5-14 Parámetros de la capa física del terminal

Categoría	Enlace descendente				Enlace ascendente	
	Número máximo de bits recibidos por TTI	Número máximo de bits transportados por un DL-SCH en un TTI	Máximo número de antenas transmisoras en el caso de multiplexado espacial	Velocidad de pico ³ (Mb/s)	Número máximo de bits recibidos por TTI	Velocidad de pico (Mb/s)
1	10296	10296	1	≅ 10	5160	≅ 5
2	51024	51024	2	≅ 51	25456	≅ 25
3	102048	75376	2	≅ 102	51024	≅ 51
4	150752	75376	2	≅ 150	51024	≅ 51
5	299552	149776	4	≅ 300	75376	≅ 75

Nota.- Recuérdese que en el caso de utilizar MIMO en el enlace descendente varios TTI pueden transmitirse en paralelo.

En el caso de multiplexado espacial, el número máximo de bits recibidos por TTI (1ms.) es la suma de los bits recibidos por los diferentes DL-SCH considerando los diferentes caminos de propagación virtuales establecidos. Por otra parte, en la tabla 5.14, la columna con el número máximo de bits transportados por un DL-SCH en un TTI indica el máximo número de bits transportados en un único DL-SCH, es decir sin considerar multiplexado espacial. En la anterior tabla se ha considerado en todos los casos que se utiliza el formato de modulación más eficiente (64 QAM) y no se han tomado en consideración mecanismos de retransmisión. Por consiguiente la velocidad de pico indicada es un valor máximo teórico, no el valor real visto por el usuario, una vez se toman en consideración las características de propagación del canal móvil y la existencia de múltiples terminales móviles accediendo al sistema.

La Tabla 5-14 también resume las características más relevantes de las diferentes categorías de terminales en el caso del enlace ascendente. Los terminales móviles de categoría 1 a 4 utilizan un formato de modulación 16 QAM mientras que los terminales móviles de categoría 5 utilizan el formato de modulación 64 QAM. Como en el caso del enlace descendente las cifras mostradas en la tabla se calculan sin tomar en consideración la influencia de los mecanismos de HARQ ni los efectos de la propagación ni del acceso multiusuario. Por consiguiente la velocidad de pico indicada es un valor máximo teórico, no el valor de la velocidad de transmisión efectiva vista por el usuario.

5.8 Procedimientos básicos asociados a la interfaz aire

5.8.1 Procedimiento de sincronización y adquisición inicial de parámetros

La primera acción que un terminal móvil LTE debe realizar después de activarse es buscar una celda a la que conectarse. Con este procedimiento se pretende:

- Sincronizarse en tiempo y frecuencia a nivel de capa física el terminal móvil con el transceptor ubicado en el eNB de la celda a la que el Terminal móvil desea conectarse.
- Adquirir la sincronización temporal a nivel de trama y subtrama
- Determinar la identidad física de la celda.

El procedimiento de búsqueda de celda se basa en el uso de las señales de sincronización primaria y secundaria del sistema. La Figura 5-24 indica las señales y canales físicos incluidos en este procedimiento.



Figura 5-24 Procedimiento de sincronización y adquisición inicial de parámetros

A partir de la señal primaria de sincronización (P-SCH), que se transmite utilizando las 62 subportadoras centrales, y utilizando un proceso de correlación, se detecta la posición temporal de la misma. En el caso de operar en modo FDD esta señal de sincronización primaria siempre aparece en el último símbolo OFDMA de la primera y undécima ranura temporal (TS) de cada trama, permitiendo de esta forma que el terminal móvil conozca los límites temporales a nivel de ranura temporal (i.e., *slot*). De modo análogo, en el caso de operar en modo TDD las señales de sincronización primarias aparecen en el tercer símbolo de las ranuras temporales tercera y décimotercera y por lo tanto, de forma análoga, la detección de estas dos señales también permite al terminal móvil fijar los límites de cada ranura temporal.

Además, con la detección del canal de sincronización primario P-SCH, el terminal móvil también realiza las siguientes funciones, [18]:

- Corrección de *offsets* de frecuencia introducidos en la señal recibida por los cabezales de RF transmisores y receptores así como el ajuste de la frecuencia central del radio-canal (*raster* en frecuencia³²).

³² El "raster" en frecuencia se define como la granularidad con que se puede ajustar la frecuencia portadora o central de un radio-canal. En el caso de LTE se ha fijado en 100KHz.

- Adquisición del sincronismo de símbolo.

A partir de la señal S-SCH, que se transmite utilizando las 62 subportadoras centrales, se detecta cual de las 168 posibles secuencias que definen la identidad de la celda es la utilizada. Como la señal S-SCH se detecta siempre después de la señal de sincronización primaria P-SCH (la señal S-SCH siempre está asociada a una señal P-SCH) es posible utilizar métodos de detección de secuencia por máxima verosimilitud (*Maximum Likelihood*, [19]), ya que se conoce la respuesta impulsional del canal que afecta a la señal S-SCH. Recordar que dicha respuesta impulsional puede obtenerse a partir de la detección de la señal de sincronización primaria. Con la detección del canal de sincronización secundario S-SCH se realizan las siguientes funciones:

- Detección del límite de trama. Para ello se toma en consideración que la señal S-SCH se transmite en el penúltimo símbolo de las ranuras temporales (i.e. *slots*) numeradas #0 y #10, correspondientes a la primera y sexta subtrama, y que los símbolos de la secuencia de 62 bits asociados a dichas señales SCH están permutados entre símbolos pares e impares (véase sección 5.5.3.1 y Figura 5-6) según se utilicen en la ranura temporal #0 o #10. Con ello se busca que una vez se detecte el contenido de la secuencia de 62 bits asociada a la señal de sincronización secundaria, automáticamente se conozca en que ranura temporal se ha detectado dicha secuencia de sincronización secundaria y por lo tanto la posición de la señal S-SCH con respecto al inicio de la trama.
- Detección del grupo de identificadores de celda (Cell group ID). En efecto, a partir de la detección de los 62 símbolos de la secuencia asociada a la señal secundaria de señalización es posible conocer a cual de los 168 posibles esquemas pertenece esta señal y por lo tanto determinar la identificación física de la celda (Cell-ID)³³.

A partir de las señales referencia o señales piloto, el terminal móvil puede estimar la respuesta impulsional del canal asociada al ancho de banda total ocupado por el radiocanal y completar la sincronización a nivel de bit, si fuera necesario. Esta información es necesaria para posteriormente poder demodular correctamente el canal de radiodifusión o *broadcast* (PBCH) así como el resto de canales de tráfico y señalización del sistema. En efecto, con la detección de las señales de sincronización primaria y secundaria efectuadas en los pasos anteriores sólo se trabaja con las 62 subportadoras centrales, por lo tanto cualquier estimación de la respuesta impulsional del canal obtenida a partir de dicha detección sólo es válida para la zona central de la canalización disponible, no para todo el ancho de banda del radiocanal. Por esta razón es necesario incluir este paso de detección de las señales de referencia, antes de proceder a detectar el canal de radiodifusión (PBCH).

Finalmente, a partir del canal de radiodifusión (PBCH), el terminal móvil obtiene información de la red (*cell system information*)³⁴ a la que quiere conectarse, información necesaria

³³ La identificación física de celda (PCI) se obtiene como producto del valor del “Cell-group ID” obtenido a partir de la señal S-SCH y del identificador de secuencia de la señal P-SCH (hay tres posibles secuencias). Por consiguiente el sistema LTE puede asignar hasta 504 PCIs distintos.

³⁴ Véase anexo 5.2 para la descripción del Sistema de Información en LTE.

para poder operar correctamente. Para ello, el terminal móvil recibe, utilizando el canal de radiodifusión (PBCH), el denominado *Master Information Block* que básicamente contiene información sobre la canalización disponible (4 bits), información sobre la configuración del canal PHICH (3 bits) y 7 bits para el número de trama (*System Frame Number-SFN*). Es importante subrayar que el terminal móvil obtiene el resto de información sobre las características e identificación del sistema a partir de los denominados *System Information Block* (SIB) que se transmiten utilizando el *Physical Downlink Shared Channel* (PDSCH). La presencia de SIBs en los canales PDSCH se indica utilizando el canal de control PDCCH.

El *System Information Block* más importante es SIB1 que incluye información sobre el operador de la celda, identidad de la celda, código del área de localización, nivel de recepción mínimo requerido, etc. En el caso de operar en modo TDD también incluye información sobre la ubicación de las subtramas ascendentes y descendente y sobre la configuración de la subtrama especial (estructura de los campos DwPTS, UpPTS y GP). También incluye información sobre la ubicación temporal (*time-scheduling*) de los restantes SIBs, que en general se agrupan en los denominados mensajes de información de sistema (*System Information* o SI). Estos SI se transmiten de forma periódica utilizando las denominadas ventanas SI (*SI-windows*) que permiten garantizar que no existe sobreposición de dichos mensajes. Véase Anexo 5.2 al final de este capítulo para más detalles sobre el Sistema de Información LTE.

5.8.2 Procedimiento de acceso aleatorio

Se entiende por acceso aleatorio al procedimiento mediante el cual un terminal móvil se conecta a un determinado eNB. Este procedimiento se ejecuta por diversas razones:

- i) al acceder el terminal móvil a la red,
- ii) cuando el terminal móvil realiza un procedimiento de *handover*, es decir cuando a lo largo de una llamada, cambia de eNB,
- iii) cuando el terminal móvil realiza procedimientos de reelección de celda,
- iv) como resultado de una llamada entrante, etc.

En el sistema LTE se definen dos tipos distintos de procedimientos de acceso:

- Acceso basado en contienda, que aplica de forma general, es decir tanto a los procesos de acceso general, transferencia de llamada (*handover*), transferencia de información en UL para establecer los mecanismos de *scheduling*, cuando no están disponibles canales del tipo PUCCH o bien cuando hay datos a transmitir en el enlace ascendente o descendente y el terminal móvil ha perdido la sincronización (e.g., debido a mecanismos de ahorro de batería o *power saving*).
- Acceso regulado (no basado en contienda), que sólo aplica en los procedimientos de *handover*.

La diferencia entre uno y otro radica que en el segundo caso no hay posibilidad de colisión de la secuencia preámbulo.

Acceso basado en contienda

El procedimiento de acceso basado en contienda sigue los pasos indicados en la Figura 5-25.

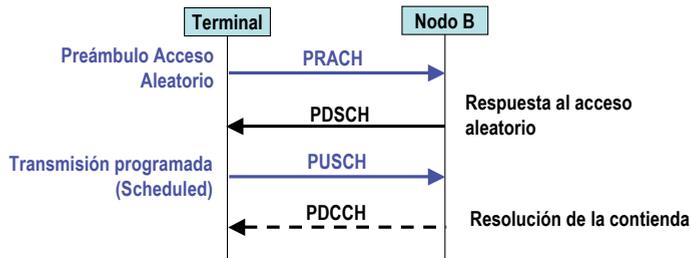


Figura 5-25 Acceso basado en contienda

1. En primer lugar se envía un preámbulo de acceso PRACH utilizando los bloques de acceso definidos a tal fin (véase Figura 5-21). La secuencia preámbulo (PRACH) se elige, de forma aleatoria, entre un grupo de hasta 64 posibles secuencias³⁵. El nivel de potencia con que el móvil envía este preámbulo se determina a partir de una estimación inicial de las pérdidas de propagación del enlace, utilizando para ello las señales de referencia (ó piloto) transmitidas en el enlace descendente.

Si no hay respuesta por parte del eNB al envío del preámbulo PRACH, al cabo de un cierto tiempo el terminal móvil repite la transmisión aumentando el nivel de potencia transmitida. A este mecanismo se le denomina *Power Ramping*.

Cabe la posibilidad de que varios terminales móviles envíen un preámbulo RACH en el mismo bloque de acceso. Si los terminales móviles han elegido secuencias preámbulo distintas, ello no es un problema ya que, dada la ortogonalidad de las secuencias de Zadoff-Chu utilizadas como secuencias preámbulo, el eNB será capaz de diferenciar entre las diferentes peticiones de acceso. En el improbable caso que dos o más terminales móviles hubieran elegido la misma secuencia preámbulo, ello tampoco sería inicialmente un problema, ya que el eNB entendería que le llega el mismo preámbulo procedente de caminos de propagación diferentes (propagación multicamino) y utilizando las propiedades del prefijo cíclico (véase capítulo 4) sería capaz de compensar dicha propagación multicamino virtual y detectar la petición de acceso.

³⁵ El número de secuencias preámbulo para acceso basado en contienda es 64 menos el número de secuencias reservado para acceso regulado. El Nodo B indica mediante mensajes de radiodifusión cuales son las secuencias reservadas para acceso regulado

2. La respuesta a la petición de acceso aleatorio contiene un duplicado del preámbulo utilizado por el móvil, que lo identifica, información para el alineamiento temporal (*time advance*)³⁶ de las siguientes transmisiones por parte del móvil, una asignación inicial del *temporary C-RNTI*³⁷ así como una asignación inicial de recursos en el enlace ascendente, necesaria para la transmisión programada que sigue a continuación.

Hay que subrayar, que en el caso de que dos o más terminales móviles hubieran utilizado la misma secuencia preámbulo en el bloque de acceso, el eNB mandaría un único mensaje de respuesta a estos terminales móviles (no diferenciaría entre ellos) ya que habría considerado que la recepción en el mismo bloque de acceso de preámbulos duplicados se correspondería con un caso de propagación multicamino del canal. En consecuencia ello obligaría a resolver este problema de colisión en los siguientes pasos del procedimiento.

3. A continuación, el terminal móvil informa al eNB, mediante un paquete PUSCH sujeto a mecanismos de retransmisión híbrida (HARQ), sobre su identidad y puede solicitar una actualización del área de seguimiento (*Tracking area*) o solicitar el establecimiento de una conexión a nivel RRC. Para ello utiliza el canal de transporte UL-SCH asignado por el Nodo B a tal fin (transmisión programada).

En el caso de que dos o más terminales móviles hubieran utilizado la misma secuencia preámbulo para el bloque de acceso en el paso 1, y en tanto que el eNB no ha diferenciado entre ellos y ha mandado un único mensaje de respuesta, todos estos terminales móviles entenderían que tienen recursos reservados para transmitir (recursos que serían los mismos) y por lo tanto se produciría una colisión entre los mensajes enviados por dichos terminales móviles. En esta situación pueden darse dos supuestos:

- El eNB no es capaz de detectar ninguno de los mensajes enviados por dichos terminales móviles, de modo que después del número de retransmisiones fijado por el mecanismo HARQ, ninguno de los terminales móviles implicados recibirá el reconocimiento positivo a la transmisión programada, y por lo tanto deberán reiniciar el mecanismo de acceso, bien inmediatamente, bien dejando pasar un tiempo elegido aleatoriamente dentro de una determinada ventana de contención, de forma similar a como se efectúa en un protocolo ALOHA.
- El eNB es capaz de detectar correctamente la transmisión de uno de los terminales móviles. En este caso el eNB enviará un reconocimiento positivo a dicho terminal móvil, mediante el canal de control PDCCH. Cuando el resto de terminales móviles implicados reciban este reconocimiento sabrán que ha habido una colisión en sus peticiones de acceso y deberán reiniciar el mecanismo de acceso tal como se ha explicado anteriormente. A este mecanismo se le denomina resolución de la contienda.

³⁶ Véase sección 5.5.4.2

³⁷ El *temporary C-RNTI* es el alias que utiliza el terminal móvil para identificarse en la red. Para evitar que se pueda suplantar a un móvil, y acceder de forma fraudulenta al sistema, la red proporciona un nuevo valor de *temporary C-RNTI* cada vez que el móvil accede al sistema.

Obsérvese como en la Figura 5-25, que describe los pasos del proceso de acceso aleatorio basado en contienda, se indica la existencia potencial de este cuarto mensaje “resolución de la contienda” (indicado mediante líneas a trazos).

Acceso regulado

El procedimiento de acceso regulado es muy similar al caso anterior, sólo que con este modo de operación no hay posibilidad de colisión ya que el terminal móvil utiliza un preámbulo de acceso previamente asignado por el eNB. El procedimiento sigue los pasos indicados en la Figura 5-26.

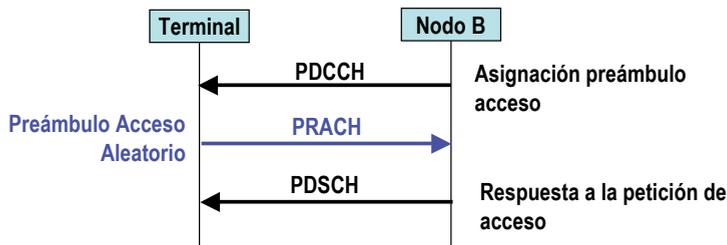


Figura 5-26 Acceso regulado

En este caso, el procedimiento lo inicia el eNB³⁸ quien asigna la secuencia preámbulo del paquete de acceso. A continuación el móvil, utilizando la secuencia preámbulo asignada, envía un paquete de acceso PRACH siguiendo el mismo procedimiento que el descrito para el acceso basado en contienda. Como respuesta el eNB envía la respuesta a la petición de acceso, tal como ocurre en el caso del acceso en contienda. No obstante ahora el proceso termina en este punto, puesto que como es la red (eNB) quien lo ha iniciado, ya conoce la identidad del terminal móvil y ha podido establecer, si fuera necesario, la correspondiente asignación de recursos para dicho terminal móvil. Ahora no es necesario considerar ningún tipo de mensaje de resolución de contienda ya que todos los mensajes intercambiados lo han sido utilizando transmisiones programadas, es decir sin competir por el acceso al canal radio.

5.8.3 Mecanismo de aviso (Paging)

El principal propósito de este procedimiento es:

1. Llamar o despertar a los terminales móviles que se encuentran en estado de espera (Idle) a nivel de capa RRC.
2. Informar a terminales móviles activos (*RRC connected*) de cambios en el sistema de

³⁸ Recordar que este procedimiento se utiliza en el procedimiento de handover, cuando éste lo inicia la red

información o de notificaciones de emergencia (ETWS) localizadas en los bloques de información de señalización SIB10 o SIB11 (véase anexo 5.2).

Cuando el terminal móvil utiliza mecanismos de recepción discontinua (DRX) para reducir el consumo de potencia, es decir opera en modo *Idle*, debe despertarse periódicamente para monitorizar el canal de control PDCCH y verificar si lleva el indicador de aviso *P-RNTI* que indica en qué recursos físicos (PRBs) del canal PDSCH se encuentra la información de aviso (i.e., *paging*). Para ello almacena un valor por defecto del denominado *paging cycle*, es decir de periodicidad del aviso y el número de grupo de aviso (i.e., *paging group*), valores transmitidos por el canal de radiodifusión (BCCH), y calcula los números de las tramas y subtramas en donde debe despertarse y acceder a la información de aviso o *paging*.

5.8.4 Mecanismo de establecimiento del servicio portador radio

Para poder establecer cualquier tipo de transferencia de información entre el terminal móvil y el eNB es necesario establecer el denominado servicio portador radio, definido en sección 2.5.2.2 del capítulo 2, tanto a nivel de señalización (servicio portador radio de señalización o SRBs) como de datos (DRBs).

El mecanismo de establecimiento de la conexión a nivel de capa RRC implica el establecimiento de un servicio portador de señalización del tipo SRB1³⁹ (dedicado al intercambio de mensajes del protocolo RRC y los protocolos NAS) y el envío de un mensaje hacia el NAS pidiendo el establecimiento de una conexión a través de la interfaz S1 que conecta al eNB con el gestor de movilidad (MME)⁴⁰.

Habitualmente, después de esta primera fase de establecimiento de la conexión se inicia una segunda fase en donde se activan los mecanismos de seguridad y se establece un segundo servicio portador radio de señalización (SRB2) y uno o varios servicios portadores de datos (DRB) dependiendo del número de portadores opcionales y por defecto establecidas con la red troncal EPC (Evolved Packet Core) del sistema LTE.

En la Figura 5-27 se describe gráficamente el procedimiento de establecimiento de la conexión y de los portadores radio.

³⁹ Véase sección 2.5 para descripción de los distintos tipos de portadora radio establecidos en el estándar LTE

⁴⁰ Véase capítulo 2 para la definición de la arquitectura e interfaces del sistema LTE.

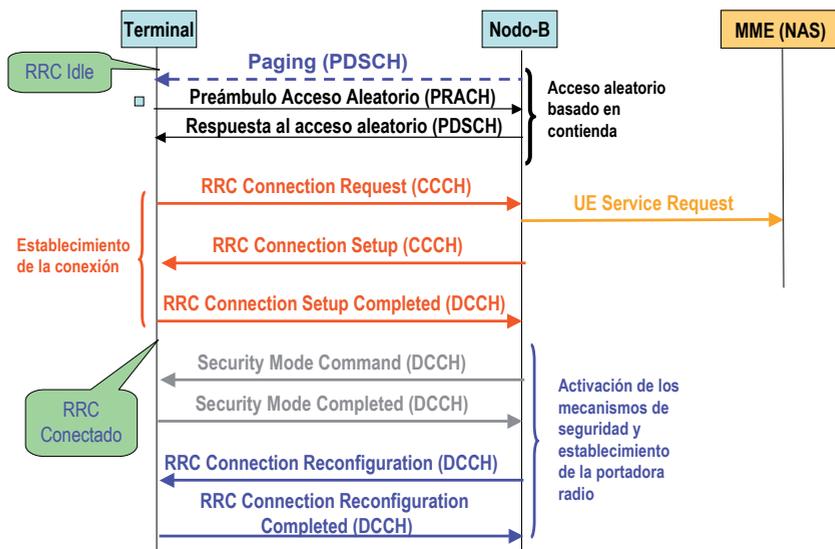


Figura 5-27 Secuencia de mensajes en el procedimiento de establecimiento de la conexión

En la Figura 5-27 se considera que el procedimiento se inicia como respuesta a un mensaje de aviso (e.g., *paging*) por parte de la red. Como consecuencia de ello el terminal móvil inicia un procedimiento de acceso aleatorio basado en contienda y posteriormente envía un mensaje de *RRC Connection Request* que incluye, además del motivo de la petición, una identificación temporal del móvil (S-TMSI) o en su defecto un número aleatorio de 40 bits. Debido a la limitada capacidad del mensaje de petición de conexión únicamente se permite identificar cinco causas de establecimiento de la conexión: Emergencia, Acceso de alta prioridad, acceso terminado en el móvil, señalización originada en el móvil, o transmisión de datos originada en el móvil.

Si el eNB acepta la petición de conexión, envía como respuesta un mensaje del tipo *RRC Connection Setup* que incluye los parámetros de configuración inicial⁴¹ de la portadora radio SRB1. Al recibir el mensaje anterior, el terminal móvil envía un mensaje de reconocimiento (*RRC Connection Setup Completed*) e incluye un mensaje al NAS, un identificador (PLMN) del operador móvil seleccionado⁴² y, si se conoce, el identificador del nodo MME al que está registrado. A partir de estos dos últimos parámetros el eNB decide cual es el nodo del núcleo de red al que debe conectarse utilizando la interfaz S1.

Una vez completada esta fase el terminal móvil cambia del estado de RRC desocupado (*RRC-Idle*) al estado RRC conectado (*RRC-connected*).

A continuación se pasa a la fase de activación de los mecanismos de seguridad (véase

⁴¹ Existe la posibilidad de que el nodo-B en lugar de dar valores individuales a todos los parámetros de configuración simplemente indique al terminal que adopte la configuración por defecto indicada en las especificaciones de la capa RRC.

⁴² En el caso de compartir la red por parte de diferentes operadores. Hay que recordar que el canal de radiodifusión (*broad-cast*) puede difundir hasta 6 identidades diferentes de redes móviles

anexo 5.5) y al establecimiento de los servicios portadores radio. Para ello en primer lugar el eNB envía un mensaje de activación del modo de seguridad (*Security Mode Command*) en donde se solicita la activación tanto del modo cifrado como de los mecanismos de protección de la integridad de los mensajes. En el mensaje también se indica qué algoritmos deben usarse. El terminal móvil verifica la protección de la integridad del mensaje recibido y configura las capas inferiores a RRC para poder aplicar a los mensajes subsiguientes los mecanismos de cifrado y protección de la integridad de la información. En todo caso hay que subrayar que el mecanismo de cifrado no se aplica al mensaje de respuesta (*Security Mode Completed*) enviado por el terminal móvil.

La red envía entonces un mensaje de petición de reconfiguración de la conexión RRC (*RRC-Connection Reconfiguration*) incluyendo los parámetros de configuración para establecer el segundo servicio portador radio de señalización (SRB2) y una o más servicios portadores radio de datos (DRB). Además de esta información de configuración, el mensaje también puede transportar señalización al NAS u otro tipo de información (por ejemplo parámetros para la configuración de los mecanismos de medidas).⁴³

El terminal móvil finalmente completa el procedimiento enviando un mensaje de reconfiguración completada (*RRC-Connection Reconfiguration Complete*).

Una vez establecidos los servicios portadores radio de señalización, ya se puede pasar a configurar una o varios servicios portadores radio de datos (DRB). El procedimiento a seguir es similar al descrito anteriormente, pero ahora el eNB decide cómo se van a enviar los paquetes de los servicios portadores EPS⁴⁴ a través de la interfaz aire. Por ello hay un mapeo unívoco de dichos servicios portadores con los servicios portadores radio (DRB), de la misma manera que estos servicios portadores radio DRB se mapean de forma unívoca con canales lógicos de tráfico dedicados (DTCH). A su vez los protocolos de capas inferiores se configuran de forma apropiada para transportar los datos de usuario. Así, el protocolo PDCP se configura para aplicar mecanismos de compresión de cabecera de acuerdo con el tipo de servicio al que corresponde un servicio portadora radio de datos (e.g., si se transmite Voz sobre IP, VoIP, la capa PDCP comprime las cabeceras de los paquetes IP recibidos para aumentar la eficiencia de la transmisión a través de la interfaz aire). De manera análoga, el protocolo RLC también configura su modo de operación de acuerdo con el tipo de servicio, mientras que el eNB asigna prioridades a los diferentes flujos de datos del enlace ascendente de acuerdo con los recursos radio disponibles. En el caso de servicios tolerantes al retardo el protocolo RLC también puede configurarse para implementar mecanismos de retransmisión.

Finalmente indiquemos que el procedimiento de establecimiento de una conexión a nivel RRC puede fallar por diversas razones como:

- El acceso puede ser denegado.

⁴³ Hay que subrayar que este mensaje puede ser enviado por el eNB antes de que se reciba la respuesta a la petición de entrar en modo cifrado (*Security Mode Completed*). En este caso, si la petición de activación del modo cifrado falla (*Security Mode Failure*), se debe de cancelar la conexión.

⁴⁴ Véase capítulo 3 para definición de los servicios portadores EPS.

- Se produce un procedimiento de reelección de celda durante el mecanismo de establecimiento de la conexión. En este caso el terminal móvil aborta el establecimiento de la conexión e informa a las capas superiores.
- El eNB puede rechazar temporalmente la conexión, en cuyo caso asigna al terminal móvil un determinado tiempo de espera antes de reiniciar el procedimiento.
- Las capas superiores NAS pueden abortar un procedimiento de establecimiento de la conexión, por ejemplo debido a que ciertos temporizadores ubicados en estas capas expiran.

5.8.5 Mecanismo de Información del estado del canal (CQI-Reporting)

El terminal móvil proporciona a la red tres tipos de información del canal:

- CQI- Indicador del estado del canal,
- RI – Indicador del rango,
- PMI- Identificador de la matriz de precodificación.

De estos tres indicadores, el más importante es el CQI que permite describir hasta 16 posibles esquemas de Modulación-Codificación (MCS). Por consiguiente, este indicador informa al eNB sobre los posibles parámetros a utilizar en el proceso de adaptación de enlace. De hecho con este indicador se informa al eNB de cual es el índice máximo de la configuración MCS que permite garantizar una tasa de error (BLER) en el bloque de transporte recibido inferior a 10^{-1} (10%). En otras palabras, el eNB podrá utilizar cualquier esquema de modulación codificación cuyo índice sea inferior o igual al indicado en el CQI, ya que está garantizado que la BLER soportada por el bloque de transporte siempre será inferior al 10%.

Con objeto de reducir el peso de la señalización, los informes sobre CQI se realizan en términos de sub-bandas, cuyo tamaño varía en función del modo de operación utilizado (periódico o aperiódico) y del número de recursos físicos (PRB) ubicados en la banda disponible. Esto es necesario para poder efectuar una gestión de recursos que tome en consideración la naturaleza selectiva en frecuencia del canal radio (*frequency selective packet scheduling*⁴⁵).

Cuando el terminal móvil opera en modo MIMO (véase capítulo 4), el indicador de rango (RI) se utiliza para indicar cual es el número apropiado de capas a considerar en el esquema de multiplexado espacial. Este indicador siempre se transmite asociado a uno o varios informes CQI, si bien como los valores de RI varían mas lentamente que los valores de CQI, este indicador es transmitido con menor periodicidad. El indicador RI toma valores 1 ó 2 en el caso de considerar estructuras de multiplexación espacial del tipo 2x2 mientras que puede llegar hasta 4 cuando la estructura de multiplexación es del tipo 4x4. En el caso de utilizar diversidad en transmisión en el eNB este indicador no aplica.

⁴⁵ Véase capítulo 6 apartado 6.4 sobre mecanismos de gestión de recursos radio.

El indicador PMI informa al eNB de cual es la matriz de precodificación⁴⁶ preferida, a tenor del estado del canal. El número de matrices de precodificación depende del número de antenas considerado en el eNB, del valor de RI y de las capacidades del terminal móvil. Así, para el caso de un eNB con dos antenas se puede elegir entre un total de 6 matrices posibles, mientras que en el caso de 4 antenas, el número de posibles matrices se eleva a 64. Como en el caso del RI este indicador se transmite en el caso de utilizar estructuras MIMO actuando como sistemas de multiplexación espacial.

Aunque cabría pensar que, cuando se detectan cambios significativos en el comportamiento del canal móvil, fuera el móvil quien activara los mecanismos de transferencia hacia el eNB del nuevo estado del canal (CQI), el sistema LTE está diseñado para que este proceso esté completamente controlado por el eNB. En efecto, para explotar por completo la ganancia aportada por los mecanismos de gestión de colas tipo *Downlink Frequency Packet Scheduler* (FDPS), [17], se requiere que el eNB tenga una información precisa del estado del canal cuando la requiera. En este contexto, la limitada capacidad (de 4 a 11 bits/subtrama) del canal PUCCH, que transmite información de control de la capa física en el enlace ascendente, impone una importante restricción. Por ello, y para optimizar el uso de los recursos del enlace ascendente, el estándar LTE considera dos opciones: informes periódicos e informes aperiódicos.

- Los informes periódicos representan el punto de partida para el mecanismo de información del estado del canal. Este tipo de informe generalmente utiliza las capacidades aportadas por el canal PUCCH, siendo el eNB quien configura los parámetros de repetición del informe. Este tipo de informe, y debido a la limitación de la capacidad del canal PUCCH, no puede contener ningún tipo de información que sea dependiente del comportamiento selectivo en frecuencia del canal⁴⁷. El periodo de repetición oscila entre 2 y 160 ms. Si en el momento en que el terminal móvil debe enviar el informe CQI, el mecanismo de gestión de recursos radio le asigna recursos al móvil para transmitir (es decir le asigna un canal PUSCH), entonces en lugar de utilizar el canal de control PUCCH se utiliza el canal de datos de usuario PUSCH.
- Cuando el eNB necesita de una información más detallada del estado del canal, puede en cualquier momento solicitar al terminal móvil un informe aperiódico, que se transmitirá utilizando el canal PUSCH. Este tipo de informes pueden adjuntarse a los propios datos de usuario transmitidos en el canal PUSCH, o transmitirse utilizando un canal PUSCH específicamente dedicado a este fin. En este caso, y en tanto que no hay un límite tan restrictivo en términos de capacidad del canal PUSCH (pueden llevar hasta 64 bits de información), los informes del estado del canal pueden ser mucho más detallados.

Los informes aperiódicos son prioritarios sobre los periódicos. Es decir, si el instante de transmisión de ambos informes coincide en un determinado momento, únicamente se envía el informe aperiódico.

⁴⁶ Véase capítulo 4 para detalles sobre las matrices de precodificación.

⁴⁷ Por ejemplo, no puede enviar información detallada (frecuencia a frecuencia) del estado del canal, sólo valores medios

5.9 Referencias bibliográficas

- [1] 3GPP TR 25.913, "Requeriments for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN)".
- [2] 3GPP TS 36.311 "Evolved Universal Terrestrial Radio Access (E-UTRA): Radio Resource Control (RRC) Protocol specification".
- [3] 3GPP TS 36.323; "Evolved Universal Terrestrial Radio Access (E-UTRA): Packet Data Convegence Protocol (PDCP) specification".
- [4] RFC3095, " Roboust Header Compression (ROCH): Framework and four Profiles: RTP,UDP; ESP; uncompressed" (<http://tools.ietf.org/html/rfc3095>).
- [5] 3GPP TS 36.322 "Evolved Universal Terrestrial Radio Access (E-UTRA): Radio Link Control (RLC) Protocol specification".
- [6] 3GPP TS 36.321 "Evolved Universal Terrestrial Radio Access (E-UTRA): Medio Access Control (MAC) Protocol specification".
- [7] 3GPP TS 36.321 "Evolved Universal Terrestrial Radio Access (E-UTRA): Medio Physical Layer Procedures".
- [8] 3GPP TS 36.104 "Evolved Universal Terrestrial Radio Access (E-UTRA): Base Station(BS) radio transmission and reception".
- [9] 3GPP TS 36.211; "Evolved Universal Terrestrial Radio Access (E-UTRA): Physical Channels and Modulation ".
- [10] 3GPP TS 36.212; "Evolved Universal Terrestrial Radio Access (E-UTRA): Multiplexing and Channel Coding ".
- [11] 3GPP TS 36.306; "Evolved Universal Terrestrial Radio Access (E-UTRA): User Equipment (UE) radio access capabilities".
- [12] J.D.C. Chu, "Polyphase Codes with Good Periodic Correlation Properties", IEEE Trans. on Information Theory, Vol. 18, pp 531-532, July 1972.
- [13] R. Franks, S. Zadoff and R. Heimailler, "Phase Shift Pulse Codes with Good Periodic Correlation Properties ", IEEE Trans. on Information Theory, Vol. 8, pp 381-382, October 1962.
- [14] 3GPP TS 23.246 "Multimedia Broadcast/Multicast Service (MBMS): Architecture and Functional Description".
- [15] M. Morelli, U. Mengali, " A comparison of Pilot-aided Channel Estimation Methods for OFDMA Systems", IEEE Transaction on Signal Processing, Vol. 49, pp. 3065-3073, diciembre 2001.
- [16] Jin Xinzhu, "Channel Estimation Techniques of SC-FDMA"; Master of Science Thesis. Karlstad University, 2007.

- [17] A. Pokhariyal, T. E. Kolding, P. E. Mogensen, "Performance of Downlink Frequency Domain Packet Scheduling for the UTRAN Long Term Evolution"; 17th IEEE Personal, Indoor and Mobile Radio Communications (PIMRC'06), Helsinki, Finlandia, 11-14, septiembre 2006.
- [18] Y. Tsai, G. Zhang, D. Grieco, F. Ozluturk "Cell Search in 3GPP Long Term Evolution Systems", IEEE Vehicular Technology Magazine, Vol. 2, pp. 23-29, junio 2007.
- [19] G. D. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," IEEE Trans. on Information Theory, Vol. 18, pp. 363–378, mayo 1972.
- [20] Llifermann, J; "Méthodes rapides de transformation du signal : Fourier, Walsh, Hadamard, Haar ", Editorial Masson, París 1979.
- [21] Hedayat, A. and Wallis, W. D. "Hadamard Matrices and Their Applications." Ann. Stat. N° 6, pp. 1184-1238, 1978.D.V. Sarwate, M.B.
- [22] Pursley, "Crosscorrelation Propieties of Pseudorandom and Related Sequences"; Proceedings IEEE, vol.68, pp. 593-619, mayo 1980.

Anexo 5.1. Secuencias utilizadas en la interfaz aire LTE

La interfaz aire del sistema LTE utiliza de modo profuso las denominadas secuencias de Zadoff-Chu por sus buenas propiedades de autocorrelación y correlación cruzada. Además, cuando necesita multiplexar informaciones de diversas fuentes o usuarios, también utiliza secuencias ortogonales de Walsh-Hadamard. Este anexo describe las propiedades principales de ambas secuencias, junto con las correspondientes a las secuencias de máxima longitud (m-sequences) y secuencias de Gold, y las relaciona con alguno de los usos que éstas tienen en el sistema LTE.

I.1 Secuencias de Zadoff–Chu

Las denominadas secuencias de Zadoff-Chu son secuencias de CAZAC (*Constant Amplitude Zero Autocorrelation Codes*) no binarias de amplitud unidad. Los elementos de una secuencia de Zadoff–Chu vienen dados por:

$$z_q(n) = \exp\left[-j\pi u \frac{n(n+1)}{N_{zc}}\right] \quad n = 0, 1, \dots, N_{zc} - 1$$

donde $u \in \{1, \dots, N_{zc}-1\}$ es el índice de la secuencia de Zadoff-Chu raíz y N_{zc} la longitud de la secuencia.

Las secuencias de Zadoff-Chu tienen las siguientes propiedades:

- Las secuencias Zadoff-Chu tienen una función de autocorrelación cíclica⁴⁸ nula excepto cuando están completamente alineadas. Esto es:

$$R_{zz}(\Delta) = \sum_{n=0}^{N_{zc}-1} z_q(n) z_q^*(n+\Delta) \equiv \begin{cases} 1 & \text{si } \Delta = 0 \\ 0 & \text{si } \Delta \neq 0 \end{cases}$$

En el contexto del sistema LTE, esta propiedad es altamente deseable ya que permite utilizar este tipo de secuencias en diversos procedimientos relacionados con la capa física, como por ejemplo para la correcta estimación del sincronismo temporal (*timing*), detección de la respuesta impulsional del canal, detección del preámbulo del acceso aleatorio, detección de señales piloto, etc. Por esta razón, las secuencias de Zadoff-Chu son ampliamente utilizadas en la interfaz aire del sistema LTE.

- Seleccionando N_{zc} como un número primo, resulta que el número de posibles secuencias que dan lugar al valor mínimo ($1/[N_{zc}]^{1/2}$) de correlación cruzada entre ellas es máximo e igual a $N_{zc}-1$.

No obstante, en el sistema LTE no siempre es posible considerar secuencias de Zadoff-Chu con N_{zc} primo. En este caso, en general, siempre es posible generar una secuencia

⁴⁸ Correlación entre una determinada secuencia y su versión desplazada circularmente

de Zadoff-Chu con N_{zc} no primo mediante la extensión cíclica o el truncamiento de una secuencia de Zadoff-Chu generada a partir de N_{zc} primo. Véase ejemplo en la Figura 5-28.

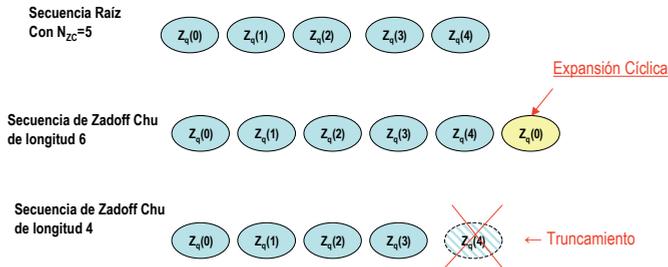


Figura 5-28.- Expansión cíclica y truncamiento de una secuencia de Zadoff-Chu

La secuencia resultante de la expansión cíclica (ó el truncamiento) de una secuencia de Zadoff-Chu conserva las propiedades de amplitud constante y autocorrelación cíclica nula, indicadas anteriormente.

- El número de secuencias de Zadoff-Chu de una cierta longitud, esto es el número de posibles valores del índice "u" en la expresión de $z_q(n)$, es igual al número de valores enteros de "u" que son primos con respecto al valor N_{zc} . Por lo tanto, con objeto de maximizar el número de secuencias raíz, normalmente se busca que N_{zc} sea un valor primo.

Como consecuencia de esta propiedad, en el caso de considerar secuencias de Zadoff-Chu de longitud corta, su número será pequeño. Esta limitación tiene relevantes implicaciones en el sistema LTE. Por ejemplo, en el enlace ascendente habitualmente se utilizan secuencias de Zadoff-Chu para implementar los denominados símbolos de referencia, que se utilizan para estimar la respuesta impulsional del canal. El número de símbolos de referencia a considerar es igual al número de subportadoras asignadas a un usuario, que siempre es un múltiplo de 12 (recordar que un PRB está compuesto por 12 subportadoras). Puesto que 12 no es un número primo, en el caso de considerar secuencias de longitud 12 ó 24 (correspondientes a la transmisión de uno o dos PRBs), el número de secuencias raíz Zadoff-Chu disponibles es muy limitado. Por esta razón, durante el proceso de estandarización de la interfaz aire LTE se buscaron (mediante simulación por ordenador) secuencias con buenas propiedades de autocorrelación y correlación cruzada y además adaptadas a la modulación QPSK. En ambos casos (1 ó 2 PRBs transmitidos) se determinó un total de 30 secuencias disponibles (ya no son secuencias de Zadoff-Chu), que están listadas en el estándar. Para un número de subportadoras igual o superior a 36 entonces sí que es posible utilizar secuencias de Zadoff-Chu. Por ejemplo, el primer número primo inferior a 36 es 31, por lo que hay 30 secuencias disponibles de longitud 31. Estas 30 secuencias sí que son utilizadas en el estándar LTE, previa extensión hasta 36 elementos utilizando un procedimiento de re-

petición cíclica. Análogamente en el caso de considerar 72 subportadoras el estándar dispone de 70 secuencias Zadoff-Chu.

- En tanto que desplazamientos cíclicos o circulares (véase Figura 5-29) de una determinada secuencia dan lugar a secuencias ortogonales entre sí, se pueden obtener múltiples secuencias con buenas propiedades de ortogonalidad a partir de una secuencia Zadoff-Chu raíz o de referencia.

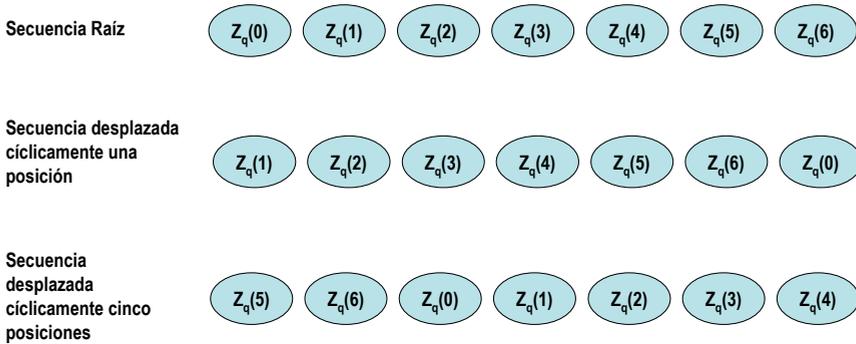


Figura 5-29.- Desplazamiento cíclico de una secuencia

Esta propiedad es ampliamente utilizada en diversos procesos relacionados con la capa física de la interfaz aire del sistema LTE. Así por ejemplo, en el caso de las secuencias preámbulo utilizadas en la transmisión del canal de acceso (PRACH) se puede partir de una secuencia raíz suficientemente larga y, mediante desplazamientos cíclicos (de tamaño suficiente para compensar los retardos por propagación multicamino introducidos por el canal de comunicaciones móviles⁴⁹) generar un conjunto de secuencias preámbulo que no se interfieran entre si.

- La transformada discreta de Fourier (DFT) de una secuencia de Zadoff-Chu es otra secuencia de Zadoff-Chu ponderada y desplazada cíclicamente. Ello implica que las secuencias de Zadoff-Chu pueden ser generadas indistintamente en el dominio del tiempo o en el dominio de la frecuencia.

Esta propiedad es relevante en el sistema LTE ya que en numerosos procedimientos asociados a su capa física se necesita operar en el dominio de la frecuencia con secuencias Zadoff-Chu, y por lo tanto no es necesario recurrir al uso de una DFT para generarlas.

1.2 Códigos de Walsh-Hadamard

Los códigos de Walsh,[20], son familias de códigos binarios ortogonales entre si. Esto es:

⁴⁹ Véase Figura 5-16

$$\overline{W(i) \bullet W(j)} \equiv \begin{cases} N & \text{si } W(i) = W(j) \\ 0 & \text{si } W(i) \neq W(j) \end{cases}$$

siendo $W(i)$ y $W(j)$ dos códigos de códigos de Walsh de la misma familia y N es la longitud de las secuencias..

Los códigos de Walsh se calculan a partir de la denominada matriz de Hadamard, [20], obtenida a partir de la siguiente ecuación recursiva:

$$H(2^k) = \begin{bmatrix} H(2^{k-1}) & H(2^{k-1}) \\ H(2^{k-1}) & -H(2^{k-1}) \end{bmatrix}$$

siendo $H(2^0) = [1]$

Por consiguiente:

$$H(2^1) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad H(2^2) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

$$H(2^3) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix}$$

Estas matrices de Hadamard tienen la propiedad de que la correlación entre dos filas (o columnas) de una misma matriz es nula. En consecuencia los códigos ortogonales de Walsh-Hadamard son simplemente las filas de la matriz de Hadamard. Es decir, denominando $W(k,N)$ al código k -ésimo de longitud N , donde $0 \leq k \leq N-1$ resulta:

$$\text{Códigos de Walsh -Hadamard de longitud 2} \rightarrow \begin{cases} W(0,2) = +1, +1 \\ W(1,2) = +1, -1 \end{cases}$$

Códigos de Walsh –Hadamard de longitud 4 →

$$\begin{cases} W(0,4) = +1, +1, +1, +1 \\ W(1,4) = +1, -1, +1, -1 \\ W(2,4) = +1, +1, -1, -1 \\ W(3,4) = +1, -1, -1, +1 \end{cases}$$

Códigos de Walsh –Hadamard de longitud 8:

$$\begin{cases} W(0,8) = +1, +1, +1, +1, +1, +1, +1, +1 \\ W(1,8) = +1, -1, +1, -1, +1, -1, +1, -1 \\ W(2,8) = +1, +1, -1, -1, +1, +1, -1, -1 \\ W(3,8) = +1, -1, -1, +1, +1, -1, -1, +1 \\ W(4,8) = +1, +1, +1, +1, -1, -1, -1, -1 \\ W(5,8) = +1, -1, +1, -1, -1, +1, -1, +1 \\ W(6,8) = +1, +1, -1, -1, -1, -1, +1, +1 \\ W(7,8) = +1, -1, -1, +1, -1, +1, +1, -1 \end{cases}$$

Es importante subrayar que los códigos de una misma longitud así generados son ortogonales entre sí, es decir:

$$\overline{W(i, N)} \bullet W(j, N) \equiv \begin{cases} N & \text{si } W(i, N) = W(j, N) \\ 0 & \text{si } W(i, N) \neq W(j, N) \end{cases}$$

1.3 Secuencias de máxima longitud (m-sequences)

Las secuencias de máxima longitud o *m-sequences* son un tipo de secuencias binaria pseudo aleatorias que se generan a partir de una función de realimentación aplicada a un registro de desplazamiento de longitud m, tal como se muestra en la Figura 5-30. La función de realimentación se obtiene como suma módulo 2 de las salidas de ciertas etapas del registro de desplazamiento.

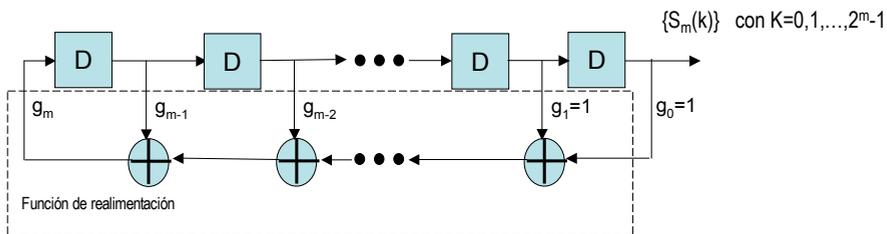
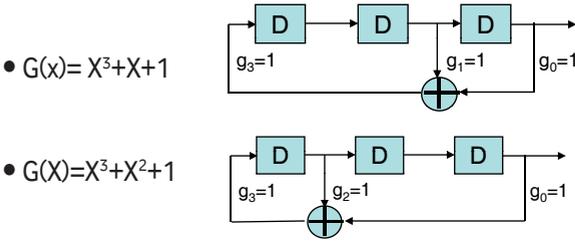


Figura 5-30 Generador de secuencia de máxima longitud

Habitualmente esta función se especifica mediante un polinomio de grado m (el número

de etapas del registro de desplazamiento) en donde los coeficientes (que pueden tomar valores de 1 ó 0) indican si la salida de una determinada etapa del registro de desplazamiento contribuye o no a la función de realimentación.

Dado un tamaño de longitud del registro de desplazamiento m , existen uno o varios polinomios que dan lugar a secuencias de máxima longitud, esto es, secuencias de longitud 2^m-1 . A este tipo de polinomio se le denomina polinomio primitivo. Por ejemplo para m igual a 3 existen dos polinomios primitivos que dan lugar a dos secuencias de longitud 7. Estos polinomios son:



En la siguiente tabla se listan algunos de los posibles polinomios generadores

m	Posibles Polinomios Primitivos
2	$X^2 + X + 1$
3	$X^3 + X^2 + 1$ ó $X^3 + X + 1$
4	$X^4 + X^3 + 1$ ó $X^4 + X + 1$
5	$X^5 + X^2 + 1$; $X^5 + X^3 + 1$; $X^5 + X^4 + X^3 + X^2 + 1$; $X^5 + X^4 + X^3 + X + 1$; $X^5 + X^3 + X^2 + X + 1$; $X^5 + X^4 + X^2 + X + 1$

Las secuencias m tienen las siguientes propiedades:

- Sea $\{X_i\}_{i \in \{1,2,\dots,2^m-1\}}$, una secuencia m , su función de autocorrelación viene dada por:

$$R_{XX}(l) = \frac{1}{2^{m-1}} \sum_{j=1}^{2^m-1} X_j \cdot X_{j-l} = \begin{cases} 1 & \text{si } l = 0 \\ -(2^m - 1)^{-1} & \text{si } l \neq 0 \end{cases}$$

Si m es suficientemente grande, se puede concluir que la función de autocorrelación de las secuencias m se aproxima a una función de autocorrelación ideal caracterizada por una delta de Kronecker ($\delta(0)$)

- Si bien no es posible obtener una expresión cerrada para la función de correlación cruzada entre dos secuencias m , $\{X_n\}$ e $\{Y_n\}$, sí que es factible proporcionar las siguiente cota,[21]:

$$|R_{X,Y}(l)| \equiv \left| \frac{1}{2^m - 1} \sum_{n=1}^{2^m - 1} X_n \cdot Y_{n-l} \right| \geq \frac{2^{(m+1)/2} - 1}{2^m - 1}$$

- Cualquier secuencia m contiene $2^{(m-1)}$ valores "1" y $2^{(m-1)}-1$ valores "0". Es decir, el número de valores "1" y valores "0" difieren en uno.
- La suma módulo 2 de una secuencia m con ella misma desplazada cíclicamente es otra secuencia m
- Si definimos como segmento de longitud "r" la secuencia de "r" símbolos idénticos consecutivos, entonces en una familia de secuencias m se tiene:
 - Un segmento de longitud m todo "1".
 - Un segmento de longitud $(m-1)$ todo "0".
 - Un segmento todo "1" y otro todo "0" de longitud $(m-2)$.
 - Dos segmentos todo "1" y otros dos todo "0" de longitud $(m-3)$.
 - Cuatro segmentos todo "1" y otros cuatro todo "0" de longitud $(m-4)$.
 - ...
 - $(2m-3)$ segmentos todo "1" y $(2m-3)$ segmentos todo "0" de longitud 1.

14.- Códigos de Gold

Las secuencias de Gold son un tipo de secuencias binarias que presentan buenas propiedades de autocorrelación, y que se generan a partir de dos secuencias de máxima longitud (*m-sequences*), denominadas secuencias preferidas, elegidas de tal modo que su función de correlación cruzada sólo tiene tres posibles valores:

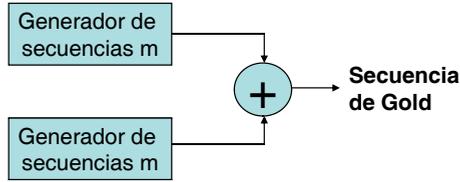
$$R_{g^{(i)}g^{(j)}} = \begin{cases} -1 \\ t(m) - 2 = R_{\max} \\ -t(m) = R_{\min} \end{cases}$$

donde

$$t(m) = \begin{cases} 2^{(m+1)/2} + 1 & \text{para } m \text{ impar} \\ 2^{(m+2)/2} + 1 & \text{para } m \text{ par} \end{cases}$$

siendo m la longitud del registro de desplazamiento utilizado para generar las secuencias preferidas

Para generar una secuencia de Gold se utiliza el siguiente esquema



Las secuencias de Gold tienen las siguientes propiedades:

- Una familia de códigos de Gold consiste en el conjunto de $2^m + 1$ secuencias, cada una de ellas de longitud $2^m - 1$. En efecto dado un par de secuencia preferidas $\{a_i\}$ y $\{b_i\}$ de periodo $N=2^m - 1$, el conjunto de códigos de Gold de dicha familia se obtiene como:

$$\begin{aligned}
 g_i^0 &= \{a_i\} \\
 g_i^1 &= \{b_i\} \\
 g_i^k &= \{a_i \oplus b_{i-k}\} \quad \text{con } k = 0, 1, \dots, N - 1
 \end{aligned}$$

siendo \oplus la función OR-exclusiva o producto base 2.

- El producto base 2 de dos funciones de Gold de la misma familia es otra función de Gold.
- Dada una familia de códigos de Gold, aproximadamente la mitad son balanceados, es decir el número de valores +1 y valores -1 difiere sólo en uno.

En la siguiente tabla se muestra el valor de la correlación cruzada máxima para distintos tamaños de la longitud del registro de desplazamiento con el que se generan los códigos de Gold.

Tamaño registro desplazamiento	Longitud Secuencia m	Valor normalizado de la correlación cruzada máxima
3	7	0,71
4	15	0,60
5	31	0,29
6	63	0,27
7	127	0,13
8	255	0,13
19	1023	0,06

Anexo 5.2.- Sistema de Información en el interfaz aire LTE

Se denomina Sistema de Información (*Information System*) al conjunto de datos transmitidos desde el eNB que informan al terminal móvil sobre los detalles de la configuración de la red LTE a la que quiere acceder.

El Sistema de Información se configura a partir de los denominados Bloques de Información de Sistema o *System Information Blocks* (SIB) en terminología anglosajona. El sistema se estructura en los diferentes tipos de SIBs que se describen a continuación:

- **Master Information Block (MIB)**, transmitido con una periodicidad fija de 40 ms. utilizando el canal de radiodifusión (PBCH), aparece en las subtramas #0 de la estructura de trama. El MIB transmite parámetros fundamentales de la red para permitir el acceso inicial del terminal móvil a la misma como, por ejemplo canalización utilizada, número de antenas en transmisión, número de trama etc.
- **System Information Block type I (SIB1)**, que se transmite cada 80ms utilizando el canal de transmisión de datos compartido (PDSCH) y aparece en la subtrama #5 de la estructura de trama. Este SIB transmite información relacionada con el acceso a la celda e información sobre asignación de recursos, como por ejemplo la identidad del operador móvil (*PLMN identity*), el área de localización, la identidad de la celda, mínimo nivel de recepción requerido, ubicación (*scheduling*) de los restantes SIBs, o en el caso de operar en modo TDD la configuración de la trama (ubicación de las subtramas del enlace ascendente y descendente), etc.
- **System Information Block type II (SIB2)**, se transmite con periodicidad de 160 ms y se utiliza para informar al móvil de la configuración de los canales comunes y compartidos, parámetros de acceso aleatorio, ancho de banda, parámetros del control de potencia en el enlace ascendente etc.
- **System Information Block type III (SIB3)**, se transmite con periodicidad de 320 ms y se utiliza para informar al móvil de parámetros comunes relacionados con los mecanismos de reelección de celda, bien sea entre celdas que utilizan el mismo conjunto de subportadoras (*intra-frequency reselection*), bien sea entre celdas que utilizan distintas subportadoras (*inter-frequency reselection*). La ubicación de este bloque de información no está predeterminada y el móvil la conoce al detectar SIB1.
- **System Information Block type IV (SIB4)**, se transmite con periodicidad de 320 ms y se utiliza para informar al móvil de parámetros relacionados con la configuración de las celdas vecinas que utilizan las mismas subportadoras (*intra-frequency cell reselection information*). Esta información es necesaria para que el móvil posteriormente pueda realizar medidas de soporte a los procedimientos de transferencia de llamada (*handover*). Como en el caso anterior tampoco está predeterminada la ubicación de este bloque de información.
- **System Information Block type V (SIB5)**, se transmite con periodicidad de 640 ms

y como en el caso anterior se utiliza para informar al móvil de parámetros relacionados con la configuración de las celdas vecinas que utilizan subportadoras distintas (*inter-frequency cell reselection information*). Nuevamente esta información se utiliza para que el móvil pueda realizar medidas de soporte a los procedimientos de transferencia de llamada. Como en los casos anteriores tampoco está predeterminada la ubicación de este bloque de información.

- **System Information Block type VI (SIB6)**, se transmite con periodicidad de 640 ms y se utiliza para dar soporte a los mecanismos de reelección hacia celdas del sistema UTRAN.
- **System Information Block type VII (SIB7)**, se transmite con periodicidad de 640 ms y se utiliza para dar soporte a los mecanismos de reelección hacia celdas del sistema GERAN.
- **System Information Block type VIII (SIB8)**, se transmite con periodicidad de 640 ms y se utiliza para dar soporte a los mecanismos de reelección hacia celdas del sistema CDMA2000.
- **System Information Block type IX (SIB9)**, se transmite con periodicidad de 640 ms y contiene el identificador del eNB (*eNB identifier*) formado por un máximo de 48 octetos. Permite al terminal móvil la opción de hacer una selección manual del eNB al que quiere conectarse.
- **System Information Block type X (SIB10)**. Contiene notificaciones del denominado *Earthquake and Tsunami Warning System*⁵⁰ (ETWS).
- **System Information Block type XI (SIB11)**. Contiene notificaciones complementarias (secundarias) del ETWS.

Como en los casos anteriores tampoco está predeterminada la ubicación de los SIB6, SIB7, SIB8, SIB9, SIB10 e SIB11, siendo el SIB1 el bloque de información que informa al móvil de la ubicación de éstos.

Los diferentes SIBs se mapean en los denominados SIs (*System Information messages*) que se transmiten utilizando PBRs del canal PDSCH. El SIB1 siempre se mapea en el denominado SI1 y, como se ha indicado anteriormente, se transmite de forma periódica cada 80 ms. De modo análogo el SIB2 siempre se mapea en el denominado SI2 y se transmite de forma periódica cada 160 ms. El resto de SIBs pueden multiplexarse entre ellos siempre de acuerdo con su periodicidad. Es decir, por ejemplo tal como se muestra en la Figura 5-31, SIB3, SIB4 y SIB5 pueden mapearse en un único SI de periodicidad 320 ms, SIB6 y SIB7 pueden multiplexarse juntos mientras que en este ejemplo SIB8, si existe, se transmite en otro SI diferente.

⁵⁰ Véase 3GPP TS 22.268 V9.2.1 (2009-06): “Public Warning System (PWS) Requirements (Release 9)”

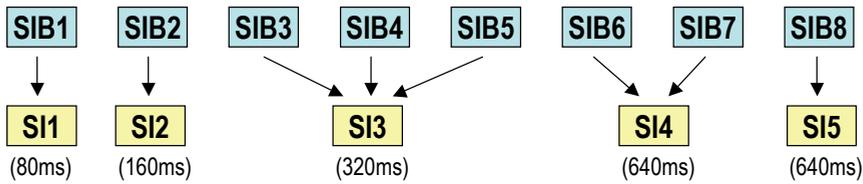


Figura 5-31.- Ejemplo de mapeo entre SIBs y SIs

Anexo 5.3.- Mecanismos de Retransmisión Híbrida (HARQ mechanisms)

El mecanismo de retransmisión híbrida se implementa a nivel de capa MAC para corregir los errores introducidos por el canal radio y garantizar una transmisión fiable a través de dicho canal.

En el modo de operación FDD el sistema LTE permite hasta 8 procesos de *Stop and Wait* – HARQ simultáneos en el enlace descendente y otros tantos en el ascendente. El mecanismo de *Stop and Wait* se basa en realizar una transmisión en el enlace correspondiente y esperar la confirmación o reconocimiento (ACK/NACK) de transmisión sin errores (ACK) o transmisión errónea (NACK) en el enlace inverso. En LTE el valor típico del tiempo de ejecución (*Round Trip Time* o *RTT*) del procedimiento de retransmisión es de 8 ms.

Cada proceso HARQ utiliza su propio *buffer* de retransmisiones y se identifica mediante un indicador (*HARQ-ID* o *HARQ process Identifier*) de 3 bits.

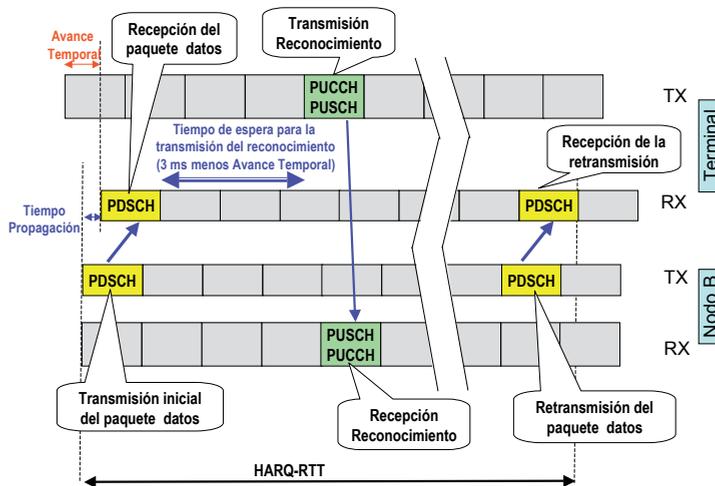


Figura 5-32 Mecanismo HARQ en el enlace descendente

En el enlace descendente, el mecanismo HARQ es asíncrono, es decir el proceso de re-

transmisión puede ocurrir en cualquier instante de tiempo después de la transmisión inicial, de modo que se necesita una señalización explícita que indique al receptor a qué paquete inicial corresponde una determinada retransmisión (ó lo que es lo mismo hay que incluir un numero de proceso HARQ para identificar las retransmisiones). En la Figura 5-32 se sintetiza gráficamente los principales pasos de este mecanismo.

Además del indicador HARQ-ID, en el enlace descendente se utilizan los siguientes indicadores:

- **NDI o New Data Indicator**.- Identifica la transmisión de un paquete nuevo.
- **RV o Redundancy Versión**.- Indica la versión de redundancia elegida para la transmisión o retransmisión (Véase apartado III.1 de este anexo).
- **MCS**.- identifica el esquema de modulación/codificación elegido. Esta selección del esquema de modulación codificación permite adaptarse a las características instantáneas del canal móvil.

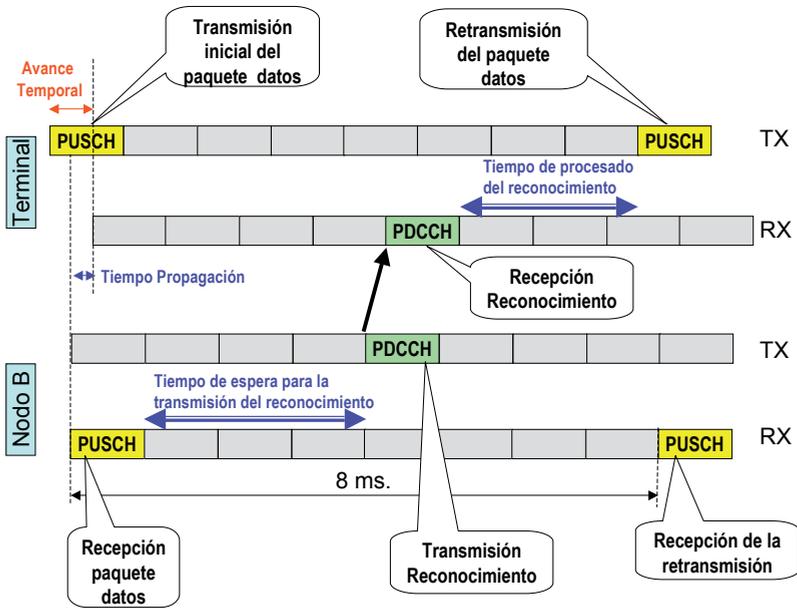


Figura 5-33.- Mecanismo HARQ en el enlace ascendente

En el enlace ascendente el mecanismo de retransmisión HARQ es síncrono, es decir los procesos de retransmisión se efectúan en instantes de tiempo predefinidos con respecto a la transmisión inicial, ello implica que no es necesario incluir un número de secuencia en el proceso HARQ. Además el proceso de retransmisión HARQ puede incluir o no mecanismos de adaptación de enlace. En el caso de no considerar mecanismos de adaptación de enlace, el mecanismo de retransmisión HARQ síncrono requiere que previamente se defina la secuen-

cia de valores RV, debidos a la ausencia de una señal de control explícita. En el caso de incluir mecanismos de adaptación de enlace el valor de RV es implícitamente indicado. En la Figura 5-33 se sintetizan gráficamente los principales pasos de este mecanismo.

En el caso de considerar un mecanismo de duplexación por división en el tiempo (TDD), si bien los fundamentos del mecanismo HARQ son básicamente los mismos, hay que considerar que las subtramas asignadas al enlace ascendente y descendente dentro de la estructura son específicas de cada celda, lo que impone ciertas restricciones tanto al mecanismo de retransmisión propiamente dicho como al envío de la información de reconocimiento. En particular, puede ocurrir que sea necesario enviar más de un mensaje de reconocimiento en la misma subtrama, tal como se muestra en la Figura 5-34. No está de más subrayar que en este ejemplo se multiplexan sobre la misma subtrama (trama 8 en la figura) dos mensajes de reconocimiento pertenecientes a dos transmisiones diferentes (M1 y M3 en la figura).

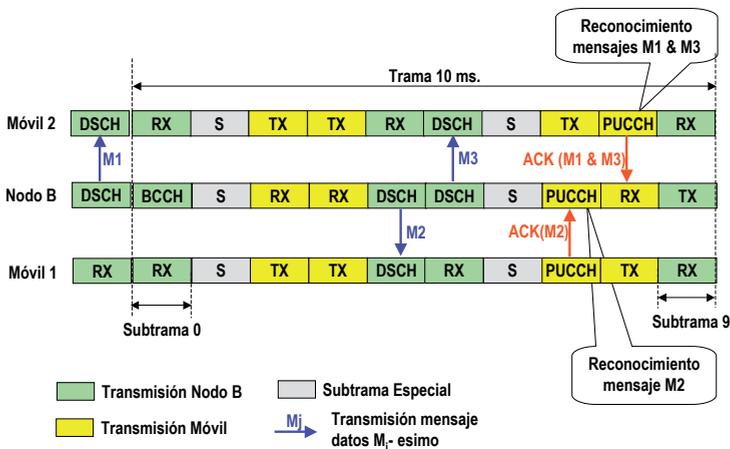


Figura 5-34 HARQ en modo TDD. Asignación de múltiples mensajes ACK/NACK a la misma subtrama

En general los valores del *Round Trip* (RTT) asociados al mecanismo de retransmisión son mayores operando en modo TDD que en modo FDD, ya que puede ocurrir que el sistema necesite esperar a que aparezcan los correspondientes recursos (subtramas) para poder enviar los mensajes. Además, también de forma genérica, el volumen de datos transmitidos por el canal PUCCH en modo TDD puede ser substancialmente mayor que el correspondiente a si se operara en modo FDD.

III.1 Redundancy Version (RV)

En LTE se utiliza un mecanismo de adaptación de tasa (*Rate Matching*) para adaptarse al número de bits que es posible transmitir en un determinado canal físico, una vez especificado su formato de transmisión. Para ello, y a partir del flujo de bits codificados a la salida del Turbo código de tasa 1/3 una vez entrelazados, se realiza un proceso de diezmado (*puncturing*) o

de repetición de bits. Además el proceso de adaptación de tasa también se utiliza en los mecanismos de retransmisión incremental (*Incremental redundancy HARQ*) para enviar nuevos bits en las posibles retransmisiones

En la Figura 5-35 se muestra un ejemplo del mecanismo de adaptación de tasa.

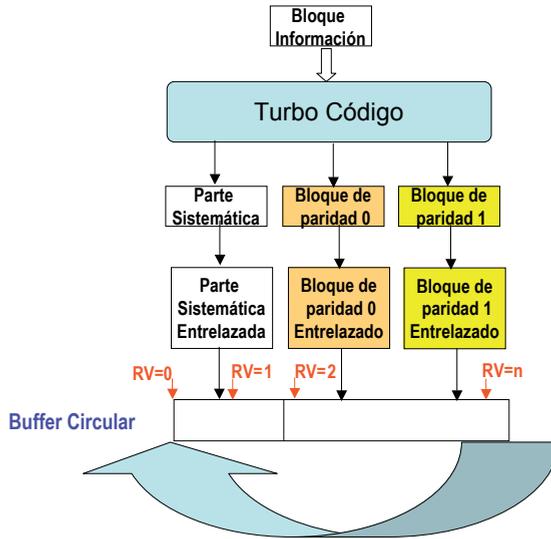


Figura 5-35.- Ejemplo de procedimiento de adaptación de tasa y selección los valores RV

En el ejemplo se considera un turbo código de tasa 1/3 que genera un bloque de bits sistemáticos (es decir idénticos a los bits del bloque de información a la entrada) y dos bloques de bits de paridad denominados 0 y 1 respectivamente. Los bits de los distintos bloques son entrelazados siguiendo para cada uno un determinado patrón. Posteriormente los bits entrelazados son ubicados en un *buffer* circular. En las primeras posiciones de este *buffer* circular se colocan los bits entrelazados procedentes del bloque sistemático, mientras que el resto de los bits (los bits de paridad entrelazados) son a su vez mezclados nuevamente y si es necesario se les aplica mecanismos de diezmo y/o redundancia.

Una vez definidos los bits de las diferentes posiciones del *buffer* circular, se establecen los valores de RV tal como se indica en la Figura 5-35. Es importante destacar que el valor del RV identifica, dentro de la estructura de *buffer* circular, el punto inicial de transmisión. Habitualmente un valor RV=0 se selecciona para la transmisión inicial. Posteriormente el gestor de paquetes e.g., *scheduler*) puede seleccionar diferentes valores de RV para la retransmisión de un paquete con objeto de optimizar los mecanismos de combinación propios del mecanismo HARQ.

Anexo 5.4. Servicio Multicast and Broadcast (MBMS)

El servicio MBMS, [14], se basa en la transmisión punto-multipunto de datos multimedia y consiste en la agrupación de dos tipos de servicios:

- Un servicio de radiodifusión o *broadcast* que puede ser recibido por todos los usuarios ubicados en el área de cobertura del eNB.
- Un servicio *multicast*, únicamente recibido por los usuarios suscritos al grupo multicast correspondiente.

Las aplicaciones típicas de este tipo de servicios pueden ser descarga de ficheros de multimedia de audio y/o video; descarga de ficheros en general; distribución de imágenes o texto; distribución de noticias, anuncios, etc. Una aplicación típica de este tipo de servicios que se prevé de gran alta relevancia en el futuro es la transmisión eficiente de señales de TV para terminales móviles, es decir adaptadas al menor tamaño de la pantalla de éstos.

Si bien el servicio MBMS puede ser soportado celda a celda de forma individual, habitualmente se piensa en este servicio para ser utilizado en un contexto de transmisión iso-frecuencia (*Single Frequency Networks*), es decir, el grupo de celdas que dan soporte a este servicio utilizan todas la misma frecuencia y están sincronizadas en tiempo para permitir que el terminal móvil trate a las múltiples señales recibidas de los diferentes eNB como si fueran múltiples ecos de una propagación multicamino, de modo que en lugar de entenderlas como interferencias las pueda combinar apropiadamente aumentando el nivel de la señal recibida. A este tipo de transmisión en la literatura se la conoce con el nombre de MBSFM (*Multicast /Broadcast over Single Frequency Networks*). Es conveniente subrayar que para permitir combinar señales procedentes de eNBs lejanos, el prefijo cíclico para este tipo de redes MBSFM tiene que ser más largo, de modo que en cada subtrama se ubican sólo 12 símbolos OFDM en lugar de los 14 habituales.

Las ventajas aportadas por la transmisión MBSFN básicamente son:

- Al permitir la combinación de las múltiples señales recibidas de los diferentes eNBs, se aumenta el nivel de la señal recibida, lo que es especialmente relevante en los límites de la zona de cobertura de las diferentes celdas.
- Se reduce el nivel de interferencia percibida por el receptor, al considerar a las múltiples señales recibidas de los diferentes eNBs como señales útiles.
- Al combinar señales con la misma información procedentes de distintos eNBs se implementa de hecho un sistema de diversidad que permite compensar los desvanecimientos introducidos por el canal radio.

En la Figura 5-36 se representa la arquitectura funcional del sistema MBSFN, en donde pueden apreciarse tres elementos funcionales nuevos:

- *Broadcast Multicast Service Center* (BM-SC). Es la entidad encargada de proporcionar los servicios a los usuarios. Es el punto de acceso a los proveedores de contenidos. Rea-

liza funciones de autorización a los terminales móviles para acceder al servicio MBMS y es el responsable de la ubicación temporal (*scheduling*) de las sesiones *Broadcast/Multicast*. También realiza funciones de confidencialidad y protección de la integridad del mensaje MBMS.

- **MBMS Gateway (MBMS GM).** Es la entidad que se encarga de distribuir los paquetes MBMS a todos los eNBs de una determinada área de servicio, así como de gestionar la sesión MBMS. También se encarga de funciones relacionadas con la tarificación del servicio MBMS.
- **Multi-cell/multicast Coordination Entity (MCE)** es la entidad lógica encargada de gestionar en frecuencia y tiempo los recursos asignados en el caso de transmisiones MBMS multicelda.

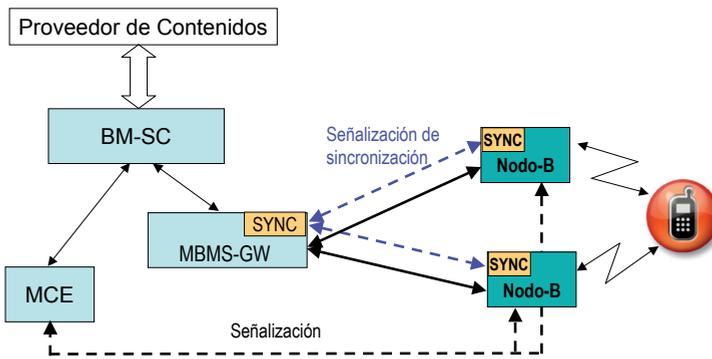


Figura 5-36. Arquitectura funcional del sistema MBSFN

Desde el punto de vista de la interfaz aire las señales MBMS se transportan utilizando el canal físico PMCH que se mapea en la estructura de trama en las denominadas subtramas MBMS (*MBMS subframes*) correspondientes al enlace descendente. En estas subtramas no se puede mapear ningún otro tipo de canal del enlace descendente.

Una subtrama MBMS, véase Figura 5-37, está formada por:

- La parte *unicast*. Esta parte es esencialmente idéntica a la parte de control de las subtramas convencionales (no MBMS) y en ella básicamente se transmiten los canales PDCCH, PCFICH y PHICH. Estos canales siempre deben de transmitirse en el enlace descendente (aunque no se transmitan en la subtrama el canal PDSCH) ya que el terminal móvil necesita información de control para poder implementar en el enlace ascendente mecanismos de reconocimiento HARQ, control de potencia y gestión de recursos (*scheduling*). Sin embargo es importante subrayar que, a diferencia de las subtramas convencionales en donde este tipo de información puede necesitar entre 1 y 3 símbolos OFDMA para ser transmitida, en el caso de la subtrama MBMS únicamente se requieren dos símbolos, como máximo. Obviamente el tamaño de esta parte debe ser el mismo para todas las celdas que componen el sistema iso-frecuencia (MBSFN).

Es conveniente destacar que en esta zona los símbolos utilizan prefijo cíclico normal.

- La parte MBMS en donde el canal de transporte MCH es mapeado. En esta parte cabe diferenciar entre los símbolos de referencia MBMS utilizados para poder estimar la respuesta impulsional del canal y poder realizar la demodulación coherente de la señal recibida, y los símbolos destinados a la transmisión de la información MBMS. Hay que subrayar que la disposición en el enrejado frecuencia/tiempo de los símbolos de referencia MBMS es específica e igual para todas las celdas de componen el sistema iso-frecuencia (MBSFN). Por otra parte también hay que destacar que la densidad de símbolos de referencia en una subtrama MBMS es substancialmente mayor que la utilizada en las subtramas convencionales (no MBMS). Ello es debido a que en una red iso-frecuencia, al recibir señales simultáneamente de diversos eNBs, la respuesta impulsional resultante percibida por el terminal móvil es más dispersiva (es como si hubieran más ecos) lo que resulta en una función de transferencia más selectiva en frecuencia y por lo tanto hay que sondear con mayor detalle el canal de comunicaciones . En esta zona los símbolos utilizan prefijo extendido.

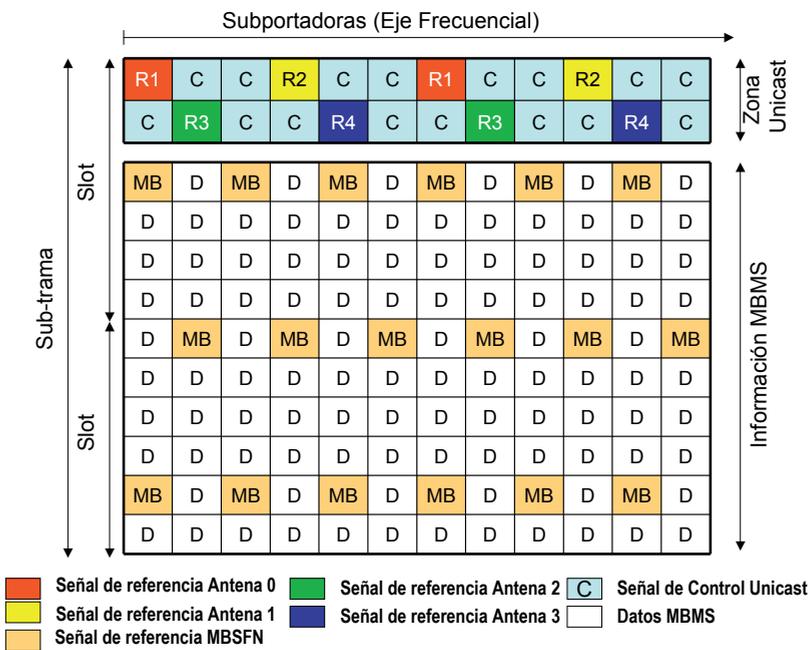


Figura 5-37 .- Estructura de una subtrama MBMS

Puesto que en una determinada subtrama MBMS sólo se puede utilizar un único patrón de señales de referencia, no existe la posibilidad de implementar ni técnicas de diversidad en transmisión ni técnicas de multiplexado espacial (MIMO).

Anexo 5.5. Mecanismos de seguridad en la interfaz aire

Los mecanismos de seguridad juegan un papel muy relevante en la interfaz aire del sistema LTE, como no podía ser de otra forma dada la naturaleza de recurso compartido de la interfaz aire. En particular se consideran dos funciones:

- Cifrado, que aplica a ambos planos (datos y señalización) de los protocolos de la interfaz aire, es decir tanto a las portadoras radio de datos (DRB) como de señalización (SRB), y garantiza que los mensajes transmitidos a través de la interfaz aire no puedan ser recibidos por terceras partes.
- Protección de la integridad, que garantiza en el plano de control que los mensajes recibidos no han sido alterados por la inserción de paquetes o la sustitución de los mismos.

El protocolo RRC activa ambas funciones simultáneamente cuando se establece una nueva conexión o bien cuando debido a un proceso de transferencia de llamada (*handover*) se establece una conexión en un nuevo eNB.

El mecanismo de seguridad de la interfaz aire del sistema LTE se basa en un sistema jerárquico de claves. En efecto, el proceso tiene como referencia una clave K_{ASME} (*Access Security Management Entity*) común y disponible únicamente en la base de datos HSS de la red y en el terminal. A partir de esta clave secreta y de un número aleatorio el Centro de Autentificación genera un conjunto de claves y valores de verificación (*checksum*) que, junto con el número aleatorio, son transferidas a la entidad de gestión de movilidad (MME). Esta entidad transmite al terminal móvil el número aleatorio y uno de los valores de verificación. Con estos valores la USIM del terminal móvil calcula su propio conjunto de claves, usando el número aleatorio recibido y la clave K_{ASME} . A continuación se procede a realizar, utilizando protocolos NAS, una autenticación mutua entre la red y el terminal móvil verificando que las claves y los valores de verificación generados en ambos extremos son los mismos.

Una vez se ha establecido la conexión a nivel RRC (véase apartado 5.8.4) el nodo MME entrega al eNB una clave de red específica denominada K_{eNB}^{51} , a partir de la que se generaran el resto de claves usadas en la interfaz aire y conocidas como *Access Stratum derived-keys*. Estas claves son:

- K_{UPenc} utilizada en los procesos de cifrado del tráfico cursado a través del plano de usuario.
- K_{RRCint} utilizada para gestionar la integridad de los mensajes de señalización generados en el RRC.
- K_{RRCenc} utilizada en los procesos de cifrado de los mensajes de señalización generados en el RRC.

Además, también se dispone de una clave intermedia, denominada *Next Hop* (NH), usada

⁵¹ A partir de la clave K_{ASME}

para generar una nueva clave, denominada K_{eNB}^* , utilizada en el proceso de traspaso de llamada entre eNBs.

Finalmente, recordar que es el protocolo PDCP quien utiliza las claves de seguridad así generadas para implementar los procesos de cifrado y protección de la integridad de la información. Es importante subrayar que estos mecanismos de seguridad nunca se desactivan, siempre que existan conexiones establecidas, aunque en el caso del cifrado es posible utilizar el modo NULL, que implica que al mensaje transmitido no se le aplica cifrado. Hay que subrayar que el algoritmo NULL también se utiliza en el caso de llamadas de emergencia realizadas sin utilizar el módulo USIM.



VI

**Gestión de Recursos Radio y del
Espectro Radioeléctrico**

6. Gestión de Recursos Radio y del Espectro Radioeléctrico

6.1 Introducción

Los sistemas de comunicaciones móviles celulares presentan un gran dinamismo debido a, entre otras causas, las variaciones de las condiciones de propagación, las variaciones de tráfico, la movilidad de los usuarios, o la interferencia. Adicionalmente, la interfaz radio ha de soportar el acceso de un número elevado de usuarios a través de servicios de distinta índole, y con unos requerimientos de calidad de servicio (*Quality of Service: QoS*) que se deben garantizar en todo momento independientemente del dinamismo del sistema. El incumplimiento por parte del operador de los requisitos de QoS podría provocar el descontento de sus abonados y por lo tanto una pérdida clara de negocio.

En un intento por respetar la calidad de servicio, el operador podría optar por sobredimensionar el número de recursos radio disponibles. Sin embargo, esta estrategia no resulta adecuada en el ámbito de las comunicaciones móviles debido a que los recursos radio son, en general, escasos y caros. En su lugar, se debe gestionar dinámicamente la interfaz radio incorporando a la red acceso un conjunto de funciones que permitan controlar y gestionar de manera eficiente los recursos radio con el mínimo sobredimensionado posible. De manera general, a ese conjunto de funciones se le denomina "Gestión de Recursos Radio" o "*Radio Resource Management*" (RRM) [1].

La gestión de recursos radio se encarga, por un lado, de racionalizar el uso de los recursos radio como frecuencias, ranuras temporales (*time slots*), potencia, o códigos según la tecnología de acceso radio, al mismo tiempo que se permite el acceso al mayor número de usuarios posible. Por otro lado, se responsabiliza de garantizar unos determinados niveles mínimos de calidad de servicio, asociados con cada uno de los servicios establecidos en el segmento radio, como paso intermedio para conseguir un cierto nivel de calidad de servicio extremo a extremo. La consecución simultánea de estos objetivos no es sencilla, ya que, el uso eficiente de recursos, la maximización de la capacidad, y el logro de la calidad de servicio suele ir en direcciones opuestas [1]. Es por ello que la gestión de recursos radio no es una función única, sino que se compone de varias funciones que se encargan de gestionar distintos aspectos de la interfaz radio, y cuya ejecución, permite en conjunto, la consecución de los distintos objetivos.

Este capítulo aborda la descripción de las funciones RRM identificadas para LTE. En concreto, el estándar LTE [2] define una serie de funciones y conceptos RRM y especifica las interfaces y señales necesarias para la ejecución de dichas funciones, pero no especifica los algoritmos o estrategias concretas que implementan dichas funciones. De esta forma, el fabricante puede diferenciar sus productos mediante la implementación de sus propias estrategias, y se permite una cierta flexibilidad para el operador a la hora de elegir los algoritmos RRM específicos que quiere para gestionar su red. Como resultado, se ha generado una gran actividad investigadora que ha dado lugar a numerosas propuestas de estrategias RRM en la literatura científica, de las cuales se ha intentado recoger en este capítulo un buen número de ellas.

En los apartados 6.2 al 6.9 del capítulo se describirán las funciones RRM contempladas por LTE, las cuales se introducen brevemente a continuación:

- Control de admisión radio (apartado 6.2).

La función de control de admisión radio se encarga de gestionar el acceso al sistema de nuevas peticiones de servicio de forma que, en caso de aceptación, se mantenga la QoS de los servicios portadores radio establecidos. Si esto no se cumpliera, entonces se rechazaría la nueva petición de servicio.

- Control de servicios portadores radio (apartado 6.3).

El control de servicios portadores radio se encarga del establecimiento y liberación de los servicios portadores radio. Además, estas estrategias se encargan del mantenimiento del servicio incluso en condiciones desfavorables, por ejemplo mediante la implementación de mecanismos de control de congestión.

- *Scheduling* de paquetes (apartado 6.4).

La funcionalidad de *scheduling* de paquetes se encarga, celda a celda, de asignar de forma dinámica recursos radio a los usuarios para que éstos puedan realizar sus transmisiones ordenadamente. Es una funcionalidad muy importante, ya que LTE está orientada a transmisión de datos y todas las transmisiones del plano de usuario se realizan a través de canales compartidos. Por lo tanto, el desempeño de estas estrategias puede afectar en gran medida a la consecución de la QoS de usuario.

- *Inter-Cell Interference Coordination* (apartado 6.5).

Como se verá, la funcionalidad de *scheduling* de paquetes asegurará la ortogonalidad entre las transmisiones de usuarios en una misma celda (es decir, éstas no se interfieren mutuamente), pero no gestiona la interferencia entre transmisiones de usuarios conectados a distintas celdas. Esta interferencia intercelular podría llegar a ser especialmente notable en LTE, provocando una degradación sustancial de las prestaciones. La función *Inter-Cell Interference Coordination* (ICIC) se encarga de mitigar la interferencia intercelular mediante la coordinación de los *schedulers* en celdas adyacentes. Es decir, ICIC permite que las decisiones de los *schedulers* en cada celda se hagan teniendo en cuenta las transmisiones planificadas en celdas contiguas, tanto evitando hacer transmisiones simultáneas en el mismo recurso radio, como controlando la potencia de transmisión. Esta coordinación es posible gracias a la existencia de la interfaz X2 de señalización entre eNBs como se vio en el Capítulo 2.

- Control de la movilidad (apartado 6.6).

El objetivo principal de las estrategias de control de la movilidad es asegurar que los terminales se encuentran en todo momento en las mejores condiciones para comunicarse con el sistema, entendiendo como mejores condiciones a aquellas impuestas, principalmente por criterios de calidad de la señal recibida, pero que podrían atender

a políticas del propio operador. Estas funciones RRM gestionan la configuración de los procedimientos de movilidad, es decir, los procedimientos de reelección de celda (en estado RRC_IDLE) y de traspaso o *handover* (en estado RRC_CONNECTED) anteriormente introducidos en el Capítulo 3.

- Balanceo de carga (apartado 6.7).

El balanceo de carga o *load balancing* es un mecanismo que se encarga de gestionar distribuciones no uniformes de la carga de tráfico entre celdas con el objetivo de ajustar la demanda de recursos radio a la disponibilidad de los mismos. Para ello, se ajustan determinados parámetros asociados a los procedimientos de movilidad para conseguir que los usuarios cambien de celda servidora y así balancear la carga soportada por cada celda.

- Control de potencia (apartado 6.8).

El objetivo del control de potencia es gestionar adecuadamente la potencia de transmisión, tanto del eNB como del terminal, de manera que se asegure una correcta recepción de la señal en/desde todos los puntos de la zona de cobertura de un eNB. En el caso del control de potencia para los terminales se persigue minimizar la potencia transmitida, ya que de este modo se consigue alargar la vida de las baterías.

- Recepción discontinua (apartado 6.9).

La recepción discontinua es un mecanismo por el cual el terminal desactiva la circuitería de recepción durante un determinado lapso de tiempo con el fin de ahorrar energía y alargar la vida de la batería.

En LTE, todas estas funciones están localizadas en el eNB, haciendo que otras entidades del *Evolved Packet Core* (EPC) como la *Serving Gateway* (S-GW) y la *Mobility Management Entity* (MME) sean independientes de la tecnología de acceso radio. En este sentido, las funcionalidades RRM en LTE tienen una arquitectura distribuida entre las distintas estaciones base. No obstante, esto no excluye que pueda existir coordinación entre estaciones base que pueda ser explotada por alguna de las funcionalidades RRM, como se ha introducido en el caso de la funcionalidad ICIC. En lo referente a protocolos, las funcionalidades RRM en LTE involucran a los protocolos de las capas física (PHY), *Radio Link Control* (RLC), *Medium Access Control* (MAC) y *Radio Resource Control* (RRC) explicadas anteriormente en el Capítulo 5. Por ejemplo, el *scheduling* de paquetes, que opera a corto plazo, está implementado en la capa RLC/MAC. Por otro lado, funciones como el control de admisión y control de servicios portadores radio, que operan en una escala de tiempo superior a la del *scheduling*, están implementadas en la capa RRC.

Finalmente, este capítulo cubre en el apartado 6.10 una característica esencial de LTE como es la alta flexibilidad que ofrece en la gestión del espectro radioeléctrico. El espectro radioeléctrico es un recurso radio escaso y muy valioso si se tiene en cuenta el coste de las licencias de adjudicación. Por lo tanto, un uso flexible y eficiente del mismo es imprescindible.

ble. Como se verá, la interfaz radio de LTE ofrece una excelente capacidad de adaptación en aspectos relativos a la gestión del espectro como la definición de las bandas de operación, la flexibilidad en el ancho de banda usado, o la asignación dinámica de espectro entre celdas.

6.2 Control de admisión radio

El control de admisión radio (*Radio Admission Control*: RAC) en un sistema de comunicaciones móviles celulares se encarga de decidir si una nueva petición de servicio para un usuario es aceptada o no. Más concretamente, una estrategia de RAC deberá comprobar, cada vez que llega una nueva petición de servicio con unos determinados requerimientos de QoS, que se puede garantizar la QoS tanto para la nueva conexión como para las conexiones ya establecidas. En caso negativo, se deniega el servicio. Por otro lado, en caso afirmativo, se admite el nuevo servicio, lo que desencadena que se ejecuten todos los mecanismos necesarios para cursarlo (es decir, se establece un nuevo servicio portador radio o *radio bearer*: RB). En definitiva, RAC permite asegurar una alta utilización de los recursos radio (aceptando nuevas sesiones mientras existan recursos radio) y al mismo tiempo garantizar la calidad de servicio de la sesiones ya establecidas (rechazando las nuevas peticiones cuando no puedan ser establecidas).

Un correcto funcionamiento de las estrategias de RAC es de vital importancia, ya que si se acepta una petición cuando en realidad no había recursos suficientes para admitirla, se produciría una degradación del servicio para algunos usuarios, provocando al operador una disminución del grado de satisfacción de sus usuarios. Por otro lado, si se rechaza una petición cuando en realidad existe capacidad para ofrecer el servicio, el operador está perdiendo eficiencia en el uso de la red y se incrementa la tasa de bloqueo de acceso al servicio aumentando igualmente la insatisfacción de los usuarios. Por lo tanto, las estrategias de RAC afectan a la consecución de la QoS en el sistema. De hecho, tal y como se explicará a continuación, el modelo de QoS definido en LTE incorpora una serie de parámetros que podrían ser utilizados por las estrategias RAC a la hora de tomar sus decisiones.

En el Capítulo 3 se presentó dicho modelo de QoS en LTE, donde un servicio portador o EPS *bearer* (comprendiendo la comunicación extremo a extremo a lo largo de la red de acceso radio y de la red troncal), se divide entre el servicio portador radio (*radio bearer*, en lo sucesivo RB) entre el UE y el eNB, y otros servicios portadores a lo largo de la red troncal. Cada EPS *bearer* tiene asociado un conjunto de parámetros de calidad de servicio, de forma que todos los paquetes pertenecientes a un mismo EPS *bearer* tienen el mismo tratamiento desde el punto de vista de la calidad del servicio. En concreto, el perfil de calidad de servicio de un EPS *bearer* se compone de los siguientes parámetros, cuya descripción se puede encontrar en el Capítulo 3:

- *Allocation and Retention Priority* (ARP).
- *Uplink/Downlink Guaranteed Bit Rate* (GBR) y *Maximum Bit Rate* (MBR).

- *QoS Class Identifier* (QCI).
- *User Equipment- Aggregate MBR* (UE-AMBR) y *Access Point Name- Aggregate MBR* (APN-AMBR).

El parámetro ARP es un parámetro de especial utilidad para las estrategias RAC. Contiene un número entero entre 1 y 15 que determina el nivel de prioridad de un EPS *bearer*, siendo 1 el nivel de mayor prioridad. En lo que respecta al control de admisión, este parámetro se utiliza para establecer prioridades a la hora de tomar decisiones, de forma que se atiende preferentemente antes a las peticiones de servicios portadores con más prioridad.

Una vez determinada la prioridad de una petición del servicio, las estrategias RAC usan el resto de parámetros para determinar si una nueva petición de servicio debe ser aceptada o no en función de la disponibilidad de recursos radio y los parámetros de QoS solicitados. Por ejemplo, una estrategia básica de RAC podría comprobar si se cumple la siguiente condición para aceptar un nuevo EPS *bearer* en el caso de servicios GBR (véase Capítulo 3):

$$\sum_i^K GBR_i + GBR_{new} \leq R_{max} \quad (6.1)$$

donde K es el número de sesiones ya establecidas, GBR_{new} es el GBR solicitado por la nueva petición de servicio y R_{max} es el máximo *throughput* admisible en la interfaz radio. Otras propuestas mejoradas de algoritmos concretos de RAC se pueden encontrar en [3] y [4].

La decisión de RAC se realiza tanto en el enlace ascendente cuando el UE inicia una conexión RRC por primera vez como en el enlace descendente cuando se establece un EPS *bearer* desde la red. Finalmente, tras la aceptación y establecimiento de un EPS *bearer*, el parámetro ARP no afecta al control de la QoS durante la transmisión de paquetes. Este control se basará únicamente en el resto de parámetros como el QCI, GBR y MBR, etc.

6.3 Control de servicios portadores radio

La función de control de servicios portadores radio (*Radio Bearer Control*: RBC) en LTE [2] está muy ligada a la de control de admisión radio, ya que, tras una aceptación de una petición de servicio por parte de RAC, se debe establecer el RB correspondiente. Es precisamente la funcionalidad del RBC la que se encarga de dicho establecimiento considerando los requerimientos de QoS para el nuevo servicio. Además, el RBC también se encarga de la liberación de RBs en los eventos de movilidad por falta de calidad o cuando se termina una sesión.

Por otro lado, el control de servicios portadores radio va más allá del establecimiento y liberación de RBs. La función RBC incluye también el mantenimiento de la QoS asociada a los mismos mientras dure la sesión [2]. Debido a las condiciones cambiantes de tráfico, propagación, movilidad, etc., es posible que las condiciones que se dieron en un determinado

momento para la admisión favorable de un conjunto de servicios no se cumplan en todo momento. En este contexto, se debe optar por soluciones que permitan la continuidad del servicio de la mayoría de las conexiones aunque ello suponga una ligera degradación de otras. En general, este tipo de estrategias recibe el nombre de control de congestión, y podrían formar parte de la función RBC en aras de mantener la QoS del acceso radio al sistema.

En concreto, el parámetro ARP de QoS asociado a un RB estandarizado en LTE, puede ser utilizado como mecanismo para ejecutar las estrategias de control de congestión [5]. Dicho parámetro, además de especificar el nivel de prioridad de un RB a la hora de evaluar las estrategias de RAC, contiene dos campos adicionales (*flags*) denominados *pre-emption capability* y *pre-emption vulnerability* que sirven para efectuar un etiquetado de los RB como se explica a continuación. Estos campos especifican si el RB asociado podría ocupar recursos que previamente fueron asignados a otros RBs de menor prioridad (campo *pre-emption capability* activado) o si por el contrario, el RB asociado podría perder los recursos asignados para admitir en su lugar a otros servicios de mayor prioridad (campo *pre-emption vulnerability* activado). Así, en situaciones de escasez de recursos radio (por ejemplo durante un pico de tráfico), RBC podría decidir liberar los recursos de los RBs de prioridad más baja, de forma que se mantenga la calidad de servicio de los RBs con más prioridad (lógicamente a costa de incrementar la probabilidad de pérdida de paquetes o *dropping rate*). Por ejemplo, esto puede ser de utilidad en situaciones de emergencia o desastre, donde se pueden liberar recursos para llamadas urgentes a través de la liberación de los servicios portadores de menor prioridad. Otra utilidad es en los servicios de video-llamada, donde la voz y el video se pueden transmitir sobre servicios portadores de distinta prioridad (i.e., de prioridad alta en el caso de la voz y baja en el caso del video). De esta forma, en caso de congestión, se perdería el video pero existiría continuidad sobre el servicio de video-llamada a través de la voz [6].

6.4 Scheduling de paquetes

El *scheduling* de paquetes en LTE consiste en asignar de forma dinámica recursos radio a los usuarios para que éstos puedan realizar sus transmisiones ordenadamente a través de la interfaz radio LTE [2]. El *scheduling* conlleva varias subtarefas como decidir, qué usuarios van a transmitir, en qué recursos de la interfaz radio, y con qué formato de modulación y codificación. Para ello, el *scheduling* de paquetes considera los requerimientos de calidad de servicio asociados a los servicios portadores radio, la calidad del canal para cada terminal, el estado de los *buffers*, etc.

El *scheduling* está muy relacionado con las características de la interfaz radio LTE, tal y como se verá a lo largo del capítulo. La interfaz radio LTE define una rejilla tiempo-frecuencia de recursos radio (Figura 6-1), donde el mínimo recurso radio asignable a un usuario es un bloque de 180 kHz en una subtrama de 1 ms, denominado "*Physical Resource Block*" (PRB), tal y como se definió en el Capítulo 5. Esta alta granularidad de la interfaz LTE permite que el *scheduler* pueda reaccionar a las rápidas variaciones del canal radio.

En el dominio temporal, es posible explotar la llamada diversidad multiusuario, donde, debido a la existencia de varios usuarios, es más probable que alguno de ellos experimente una buena calidad del canal en una determinada subtrama. Así, si el *scheduler* tiene en cuenta el estado del canal de cada usuario (*channel-aware scheduling*), podría asignar el canal al usuario con mejor calidad y de este modo se podría aprovechar mejor los recursos radio. Esta mejora por diversidad multiusuario es tanto mayor cuanto mayor es el número de usuarios y más rápidas son las variaciones del canal.

LTE permite además extender este concepto al dominio de la frecuencia (Figura 6-2), donde ahora se tienen en cuenta las variaciones del canal debido a los desvanecimientos selectivos en frecuencia, y que son apreciables en los sistemas celulares donde el ancho de banda de coherencia¹ es típicamente menor que el ancho de banda del sistema². En la figura se representa el canal para dos usuarios. Se aprecia que tanto en frecuencia como en el tiempo el canal es percibido de manera diferente por cada usuario, de forma que si se selecciona convenientemente el usuario a transmitir, entonces es posible que el sistema trabaje más eficientemente (e.g., más próximo a su capacidad).

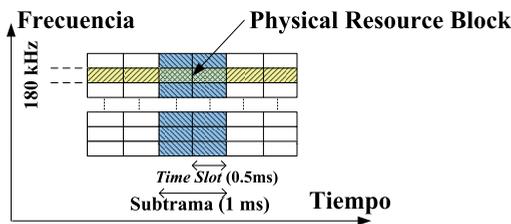


Figura 6-1 Rejilla tiempo-frecuencia en LTE

Representación del canal percibido por 2 usuarios

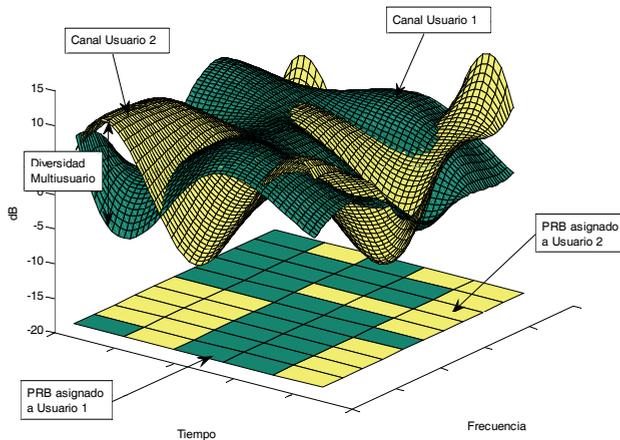


Figura 6-2 Ilustración de la diversidad multiusuario y su posible explotación mediante *scheduling*

¹ El ancho de banda de coherencia determina la anchura espectral donde la respuesta en frecuencia del canal se puede considerar plana, y por lo tanto la señal no sufriría distorsión en frecuencia..

² El modelo de canal TU (Tabla C.3.3 [7]) tiene un ancho de banda de coherencia de aproximadamente 190 KHz mientras que LTE se desplegará en una banda de al menos 1.25 MHz

En LTE el *scheduling* reside en el eNB, tanto para el DL como para el UL. Este mecanismo centralizado asegura ortogonalidad en las transmisiones, aunque requiere de señalización entre los usuarios y el eNB para reportar, entre otras cosas, el estado del canal y del buffer. Es importante destacar que, aunque la función de *scheduling* se ejecuta en cada eNB, el *scheduler* podría considerar determinadas restricciones o preferencias provenientes de otros eNB a la hora de asignar algunos de los recursos de la interfaz radio, con el objetivo de coordinar sus decisiones y mitigar la interferencia intercelular. Los mecanismos que permiten determinar dichas restricciones o preferencias entran dentro de la funcionalidad *InterCell Interference Coordination* (ICIC) que se explicará en el apartado 6.5.

La Figura 6-3 muestra la arquitectura para el scheduling en LTE, donde se distingue entre un *scheduler* para el DL (*downlink scheduler*) y otro para el UL (*uplink scheduler*).

En el enlace descendente, el *downlink scheduler* se encarga de decidir qué usuarios transmiten en el canal PDSCH, cómo se multiplexan sus canales lógicos (*MAC multiplexing*), y con qué formato de transporte (TF), que determina el tamaño del bloque de transporte (TB), y la modulación y codificación a emplear (véase el subapartado 6.4.4). Debido a las variaciones en frecuencia del canal, el *scheduler* podría asignar recursos a varios terminales en una misma subtrama. Para ello, el *scheduler* necesita que los terminales reporten el estado del canal mediante *channel status reports*. La decisión de scheduling se envía a los terminales a través del PDCCH, donde se detalla, para cada PRB, la identidad del terminal al que se le ha asignado el PRB y toda la información necesaria para que pueda decodificar el TB. Este mensaje recibe el nombre de *downlink scheduling assignment*.

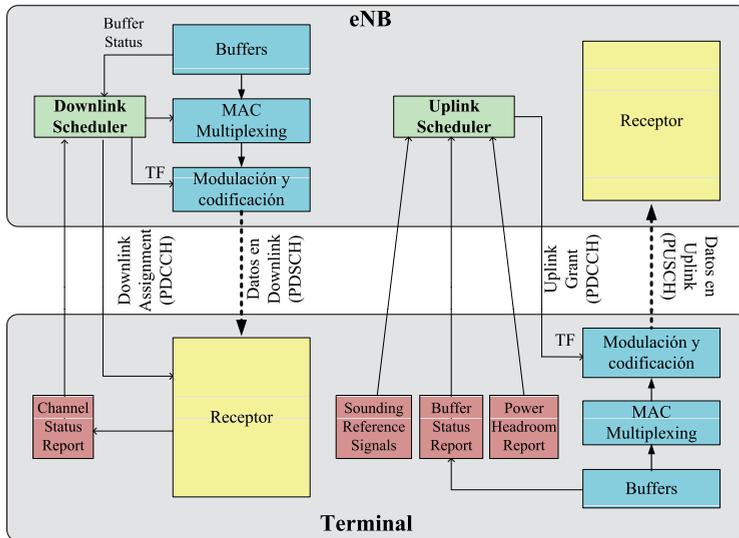


Figura 6-3 Esquema funcional para scheduling en LTE

En el enlace ascendente, de manera similar al enlace descendente, el *uplink scheduler* ubicado en el eNB determina qué terminales transmiten en el PUSCH en cada subtrama y en qué PRB. Sin embargo, debido a las características del acceso radio en UL, basado en SC-FDMA, los PRB asignados a cada usuario tienen que ser contiguos en frecuencia (véase el Capítulo 4). Adicionalmente, el *uplink scheduler* se encarga de seleccionar el TF que el terminal debería usar, mientras que las funciones de multiplexado de canales lógicos (*MAC multiplexing*) se realizan en el terminal. Tanto la decisión de scheduling como el TF seleccionado se comunican al terminal mediante los llamados *uplink scheduling grants*. Estos mensajes son válidos para una subtrama y, al igual que para las asignaciones en DL, se envían al terminal a través del PDCCH. Por lo tanto, el *uplink scheduler* conoce de antemano el TF (porque lo decide él), y no es necesario que el terminal le indique esta información al eNB al hacer sus transmisiones, lo que reduce la señalización fuera de banda en UL. Sin embargo, para seleccionar adecuadamente el TF, el eNB necesita conocer con precisión el estado del canal, el estado del *buffer*, y la disponibilidad de potencia en el terminal. Como se verá a continuación, el estándar define mecanismos específicos para que el terminal pueda proporcionar esta información.

6.4.1 Información de estado del canal

Para realizar *channel aware scheduling* es necesario que el *scheduler* tenga una estimación del canal para cada usuario tanto en DL como en UL.

Estimación de canal en DL

En lo que respecta a la estimación del canal en DL, el eNB configura a los terminales para que le envíen *channel status reports* (CSR) de todo el ancho de banda de *scheduling* o de una determinada subbanda. En concreto, el envío de CSRs por parte de los terminales se puede configurar de forma periódica o a demanda según el criterio del eNB. Los CSRs se envían a través del PUSCH si el terminal tiene un *uplink grant* o, en su defecto, a través del PUCCH. Es importante destacar que cada CSR no contiene información del estado del canal que percibe un usuario en DL, sino que contiene una recomendación por parte del terminal de los parámetros de transmisión para el DL más adecuados en función del canal percibido (típicamente la SINR). Así, el CSR puede contener la siguiente información [8]:

- *Rank Indication* (RI): Es el rango del canal para transmisión MIMO que el terminal recomienda que se use. Esta indicación sólo la deben aportar los usuarios que están configurados en modo multiplexación espacial (*spatial-multiplexing*).
- *Precoder Matrix Indication* (PMI): Recomienda la matriz de pre-codificación (Capítulo 4) que debe ser usada. Al igual que el RI, sólo se reporta esta recomendación si el terminal está configurado en modo *spatial-multiplexing*.
- *Channel Quality Indication* (CQI): Representa la modulación y codificación en base al estado del canal medido, donde el CQI indexa una tabla con una serie de esquemas de modulación y codificación predefinidos (Capítulo 5).

Una vez recibido el CSR, es el eNB el que toma la decisión final de seguir o no la recomendación hecha por el terminal. En el caso de que el eNB acepte la recomendación de RI y PMI, se confirma al terminal a través de 1 bit en el *downlink scheduling assignment*. En otro caso, se detalla explícitamente la configuración de *spatial multiplexing* usada. Por otro lado, el *downlink scheduling assignment* siempre contiene la información del TF para modulación y codificación que debe usar el terminal, y que, según la decisión tomada por el eNB, puede o no coincidir con la recomendación hecha en el CQI.

Estimación de canal en UL

En el UL, el eNB puede conseguir la estimación del canal simplemente analizando la señal recibida en los PRBs asignados a un determinado usuario para la transmisión. Sin embargo, esto requiere que el terminal esté transmitiendo (tenga *uplink grants*) y no aporta información con respecto al resto de PRBs no asignados. Para ello, el eNB puede configurar a los usuarios activos para que transmitan unas señales específicas denominadas *Sounding Reference Signals (SRS)* que permiten estimar el canal en cualquier PRB [8].

Las SRS son unas señales predefinidas y conocidas por el eNB que los usuarios transmiten de forma periódica en un determinado ancho de banda. El eNB se encarga en todo momento de decidir qué terminales activos las transmiten, durante cuánto tiempo y con qué periodicidad, comunicándoselo a los terminales a través de señalización de la capa RRC. También decide el ancho de banda que cubren las SRS. Para ello, el eNB tiene en cuenta que cuanto más amplio es el ancho de banda de SRS más consumo de potencia y menos precisión en las medidas [9]. Esto es así porque a mayor ancho de banda, más baja es la densidad de potencia resultante (para una determinada potencia máxima), y por lo tanto existe la posibilidad de degradar la estimación de canal. Por otro lado, cuanto más estrecho es el ancho de banda cubierto por las SRS, se necesitan más transmisiones en distintas partes del espectro disponible para monitorizar todos los canales (aumenta la sobrecarga de señalización u *overhead*).

Por otra parte, es posible configurar a varios usuarios para transmitir SRS en los mismos PRBs al mismo tiempo, ya que las SRS están codificadas de forma que existe ortogonalidad entre ellas (siempre que los niveles de potencia recibidos sean adecuados). Es por ello que existe un mecanismo de control de potencia específico para ellas. Para detalles adicionales acerca de las SRS véase el Capítulo 5.

6.4.2 Información de estado del buffer

Las estrategias de scheduling deben tener en cuenta el estado de los *buffers* a la hora de realizar la asignación de recursos. Por ejemplo, no tendría sentido asignar recursos en una subtrama a un terminal con un excelente canal pero que no tiene información a transmitir. En el caso del DL, el *scheduler* conoce exactamente el estado de los *buffers* en el eNB. Sin

embargo, en el caso del UL, el estándar ha definido una serie de mecanismos para que el terminal pueda reportar al *uplink scheduler* en el eNB el estado de sus *buffers*, como se describe en lo sucesivo.

Un terminal puede reportar a la red el estado de sus *buffers* en cada subtrama mediante el *Buffer Status Report* (BSR). Este mensaje se incluye en la señalización de control en la cabecera MAC durante las transmisiones en el PUSCH (señalización en banda) [10]. Un BSR indica la cantidad de datos esperando en los *buffers* del terminal para cada canal lógico. El envío de BSRs puede deber, (1) a la llegada de datos para un canal lógico de mayor prioridad que el que se está transmitiendo para que así lo tenga en cuenta el *scheduler*, (2) al cambio de celda servidora, (3) a la expiración de un contador (transmisión periódica), o (4) a la necesidad de un relleno de datos excesivo (es decir, en lugar de completar un *transport block* con bits de relleno hasta alcanzar un *transport block size*, se inserta un BSR si hay espacio suficiente para ello, optimizando así las transmisiones).

Es importante ver que esta información sólo se puede enviar si el terminal tiene recursos asignados (i.e., dispone de un *uplink grant* para la trama actual). En el caso de que el terminal no tenga recursos asignados y tenga datos en el *buffer*, necesita solicitar al eNB permiso para transmitir (un terminal nunca puede transmitir de forma unilateral ya que en ese caso no estaría asegurada la ortogonalidad en UL). Las solicitudes de transmisión se pueden realizar de dos modos en LTE: mediante *Scheduling Request* (SR) indicando que el terminal tiene nuevos datos preparados para enviar o a través del RACH (Capítulo 5).

El SR se envía a través del PUCCH y consiste en un bit dedicado a cada terminal indicando que el terminal tiene nuevos datos en el *buffer*. Para que un terminal pueda tener reservado este bit en el PUCCH necesita estar sincronizado y alineado temporalmente con el eNB, habiendo transmitido en UL en al menos los últimos 100 ms. Aún en este caso, el eNB puede no reservar este bit en el PUCCH debido a limitaciones de capacidad. La disponibilidad o no del bit se le comunica al terminal a través de señalización de la capa RRC. Siempre que el terminal quiera transmitir y no encuentre un bit dedicado en el PUCCH deberá notificar a la red que quiere asignación recursos en UL a través del RACH.

La ventaja de SR frente a RACH es que SR es un mecanismo de acceso al medio libre de contienda, donde no puede haber colisiones. Por lo tanto, es un mecanismo más rápido para obtener asignación de recursos en UL. Una vez que el terminal tiene *uplink grant* puede enviar los BSRs en sus transmisiones.

6.4.3 Información de disponibilidad de potencia

En LTE, se define la disponibilidad de potencia como la diferencia entre la potencia máxima en el transmisor y la potencia nominal de transmisión necesaria para asegurar la correcta recepción de los datos con un determinado formato de modulación y codificación. Como se verá en el siguiente apartado, cada formato de modulación y codificación funciona correctamente bajo unas condiciones mínimas de calidad de la señal recibida, que a su vez depende

de la potencia transmitida. Por lo tanto, a la hora de hacer el *scheduling*, el *scheduler* (DL o UL) en el eNB debe conocer la disponibilidad de potencia en el transmisor para seleccionar adecuadamente el formato de modulación y codificación.

En lo referente al DL, la información de disponibilidad de potencia es conocida por el *downlink scheduler* ya que éste reside en el mismo eNB donde se encuentra el transmisor, y la potencia total de transmisión es típicamente constante.

En cuanto al UL, se hace necesaria la comunicación entre el terminal y el eNB donde reside el *uplink scheduler*, con lo que LTE estandariza un mecanismo por el cual el terminal puede informar al eNB de su disponibilidad de potencia en un mensaje denominado *power headroom* [8]. Al igual que los mensajes de estado del *buffer*, este mensaje se encuentra en la cabecera MAC en las transmisiones a través del PUSCH [10]. Este mensaje se envía siempre que expire un determinado contador periódico, o cada vez que se produzca un cambio importante en el *pathloss* estimado (con una mínima separación entre mensajes para evitar notificaciones excesivamente sucesivas).

En lo que respecta al valor reportado en el *power headroom* en UL, como se verá en el apartado 6.8, su cálculo es más complejo que en el DL, ya que en UL se implementa control de potencia y por lo tanto, depende indirectamente de la distancia entre el terminal y el eNB, y de la interferencia. Como se verá en el apartado 6.8, el *power headroom* podría llegar a ser negativo, indicando que el terminal está limitado en potencia. En ese caso, el *uplink scheduler* deberá seleccionar, para las transmisiones del terminal, los TF con modulaciones más bajas ya que requieren una menor densidad espectral para su correcta recepción.

6.4.4 Adaptación dinámica del enlace

Una de las subtareas del *scheduler* en LTE es la selección del formato de modulación y codificación más adecuado en todo momento para las transmisiones de cada usuario. A este proceso de decisión se le denomina "adaptación dinámica del enlace". En LTE se implementa un mecanismo de adaptación dinámica del enlace basado en control de la tasa binaria (*rate control*), que consiste en ir variando la tasa binaria asignada a un usuario activo en función de la calidad del canal. Así, el usuario percibirá una tasa binaria alta si la calidad del canal es buena, y baja si la calidad del canal es mala.

Básicamente, este procedimiento está esquematizado de forma general en la Figura 6-4. La adaptación dinámica del enlace se basa en la estimación de la calidad del canal que, para el UL, se obtiene a partir de la recepción de las *Sounding Reference Signals* (SRS) y, para el DL, viene dada en la información contenida en los *Channel Status Reports* (CSR). A partir de esta estimación es posible determinar un esquema de modulación y codificación para garantizar una determinada *Bit Error Rate* (BER) que dependerá del servicio ofrecido y de los esquemas de codificación disponibles. Como ejemplo, la Figura 6-5 muestra, junto al límite teórico de capacidad de Shannon, una tabla genérica para obtener una BER de 10^{-2} consiguiendo la máxima velocidad de transmisión posible para cada valor de SINR (considerando

que la estimación de la calidad del canal se pudiese traducir en distintos valores de SINR). Se puede apreciar que para SINRs bajas, se emplean mecanismos de modulación y codificación más robustos como BPSK, a costa de tener una baja eficiencia espectral y por lo tanto reducir la tasa binaria.

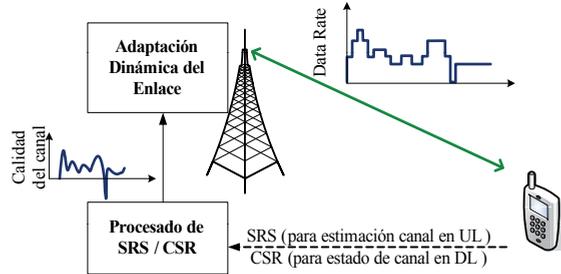


Figura 6-4. Adaptación del enlace basada en Rate Control

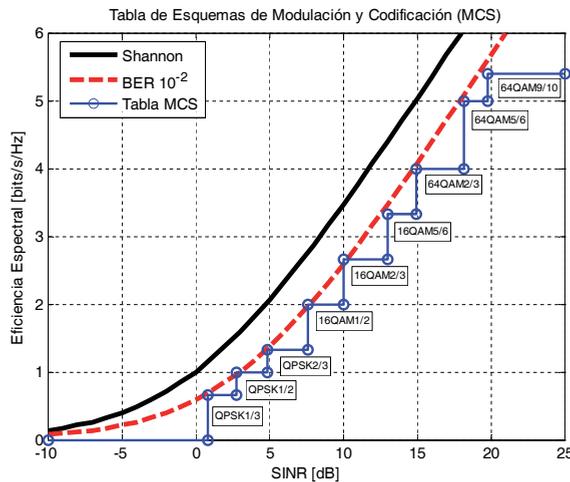


Figura 6-5. Ejemplo de tabla de modulación y codificación en función de la SINR

En LTE las transmisiones se organizan en *transport blocks* (TB). La adaptación dinámica del enlace se realiza mediante la selección del *Transport Format* (TF). El TF especifica cómo se van a realizar las transmisiones a través de la interfaz radio para un determinado *radio bearer*. En concreto se especifica el tamaño del TB, el esquema de modulación y el *mapping* en las antenas en el caso de transmisión MIMO. El *transport format* se envía a través del denominado *Downlink Control Indicator* (DCI) en el canal PDCCH. De hecho, el DCI también transporta otra información de *scheduling* para el DL o UL como el número de PRBs y PRBs asignados (véase el Capítulo 5 apartado 5.5).

El formato de modulación y codificación y el tamaño del TB se especifican con 5 bits en

el DCI. Éstos 5 bits permiten indexar una tabla de 32 entradas, donde para cada entrada se especifica una modulación y un índice a otra tabla donde se especifica el tamaño del TB (en bits) en función del número de PRB asignados [8]. La tasa de codificación resultante se puede obtener conocida la modulación, el tamaño del TB y el código de canal. Los esquemas de modulación utilizados en DL son QPSK, 16 QAM y 64 QAM, mientras que para el UL se contempla QPSK y 16 QAM y opcionalmente 64 QAM dependiendo de las capacidades del terminal.

6.4.5 Estrategias de scheduling

El estándar define los mecanismos para hacer posible el *scheduling* de paquetes, pero no define las estrategias específicas para decidir qué usuarios transmiten y en qué PRBs en una determinada subtrama. Estas estrategias serán pues, presumiblemente, de implementación propia por parte del fabricante. A continuación se describen algunas estrategias típicas. El objetivo de estas estrategias es decidir, para cada PRB n y en cada subtrama t , el usuario que debe transmitir, denotado con $m_n^*(t)$.

- **Maximum Rate (MR)** [11]: Esta estrategia selecciona al usuario que experimenta, en cada subtrama, las mejores condiciones del canal radio en un PRB dado. Matemáticamente, el criterio MR se puede expresar como:

$$m_n^*(t) = \underbrace{\arg \max}_m \{ R_{m,n}(t) \} \tag{6.2}$$

donde $R_{m,n}(t)$ es la tasa binaria de datos (*data rate*) instantánea que puede alcanzar el usuario m en el PRB n de la subtrama t . Esta tasa se obtiene mediante los procesos de adaptación del enlace, donde dada una determinada calidad del canal se puede establecer el esquema de modulación y codificación más adecuado, lo que determina una eficiencia espectral en bits/s/Hz. Por lo tanto, conocido el ancho de banda de un PRB, se puede extraer $R_{m,n}$. MR permite maximizar la capacidad del canal gracias a la diversidad multiusuario, ya que si el número de usuarios es suficientemente grande, es posible encontrar a uno de ellos que pueda transmitir en tasas cercanas a la máxima capacidad. Sin embargo, MR no es justo (*“fair”* en términos anglosajones) en algunos escenarios, donde un usuario que experimentase una calidad del canal muy superior al resto (por estar muy cerca de la estación base) obtendría casi siempre el canal, impidiendo las transmisiones del resto de usuarios, lo que es inaceptable desde el punto de vista de la calidad del servicio.

- **Round Robin (RR)** [11]: Esta estrategia asigna cíclicamente el canal a los usuarios sin considerar las condiciones del canal radio (es decir no es *channel aware*). Por lo tanto, es una estrategia justa en el sentido de que cada usuario tiene asignado el canal durante el mismo tiempo. Sin embargo, al no considerar las condiciones del canal, no está adaptado a las necesidades de los usuarios, donde, presumiblemente, aquellos con

peor condición del canal radio requerirán más tiempo el canal para igualar su calidad de servicio a la de los usuarios con las mejores condiciones radio.

- **Proportional Fair (PF)** [11]: Esta estrategia trata de combinar un uso del canal radio cercano a la máxima capacidad con un reparto justo de los recursos entre los usuarios. Para ello, se hace la selección del usuario a transmitir en base al siguiente criterio:

$$m_n^*(t) = \underbrace{\arg \max}_m \left\{ \frac{R_{m,n}(t)}{W_{m,n}(t)} \right\} \quad (6.3)$$

donde $W_{m,n}(t)$ es la tasa binaria media que ha obtenido el usuario m en un determinado periodo de tiempo. Nótese que $W_{m,n}(t)$ no es el promedio de $R_{m,n}(t)$, ya que $R_{m,n}(t)$ refleja la tasa binaria que obtendría un usuario en caso de obtener el canal, mientras que $W_{m,n}(t)$ se obtiene con la tasa finalmente asignada.

La Figura 6-6 compara estas estrategias en un escenario con tres usuarios con una SINR media de 20 dB, 10 dB y 0 dB respectivamente, donde se ha empleado un modelo de canal TU (Tabla C.3.3 [7]) y una duración de una subtrama de 1 ms. En la fila superior se observa la SINR instantánea para cada usuario y el usuario que selecciona cada estrategia de *scheduling*. Nótese como MR selecciona el usuario con la mejor SINR mientras que RR reparte equitativamente las oportunidades de transmisión. En la fila inferior vemos la tasa binaria que se obtiene en la celda y la tasa binaria media por usuario. MR claramente hace que uno de los usuarios obtenga un *data rate* superior al de los otros dos. Obsérvese que esto no es así en el caso de RR y PF.

Por último, la Figura 6-7 ilustra que para poder explotar la diversidad multiusuario es necesario que el *scheduler* sea *channel-aware*. En la figura se observa la eficiencia espectral que se obtiene en una celda en el caso de RR y PF. Esta eficiencia espectral se obtiene bajo distintos patrones de relación señal a interferencia y ruido en la celda. La máxima eficiencia espectral obtenible en este ejemplo es de 5 bits/s/Hz (correspondiente a una modulación 64 QAM con una tasa de codificación de 5/6). En primer lugar, hay que destacar que, mientras la eficiencia espectral de RR no muestra dependencia con el número de usuarios por tratarse de un *scheduler* que no tiene en cuenta el canal percibido por los usuarios, PF sí muestra tal dependencia, incrementándose la eficiencia espectral con el número de usuarios. Ciertamente, al tratarse de un *channel-aware scheduler*, PF es capaz de explotar la diversidad multiusuario. Por otro lado, a partir de 10 usuarios la ganancia de PF sobre RR se estabiliza ya que con esos usuarios es posible garantizar que algún usuario percibe las mejores condiciones del canal. Finalmente, como es lógico para un sólo usuario, ambos *schedulers* obtienen el mismo resultado.

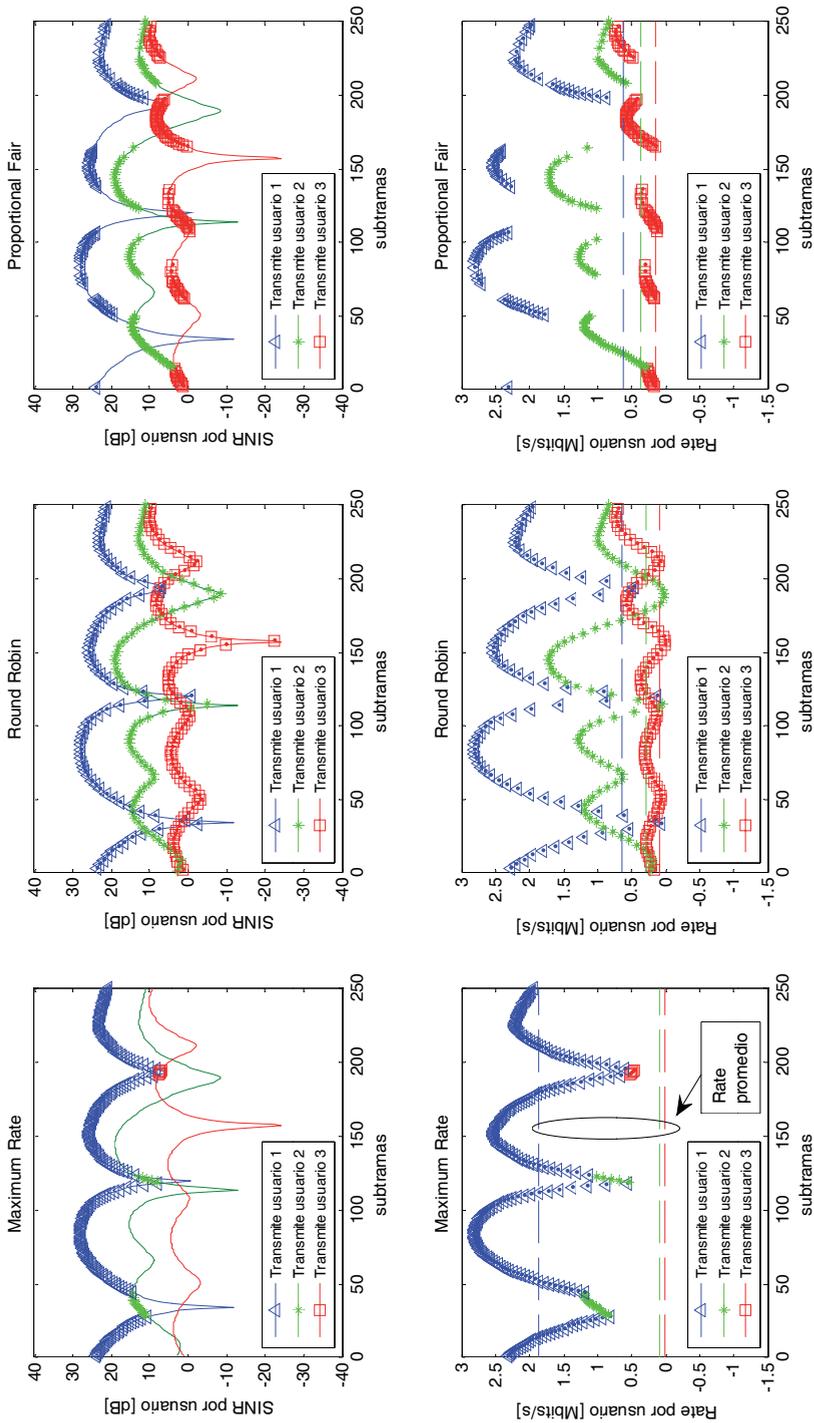


Figura 6-6 Comparación entre estrategias de *scheduling*. Arriba: SINR para tres usuarios y usuario seleccionado para cada estrategia. Abajo: Tasa binaria instantánea en la celda y tasa binaria media por usuario

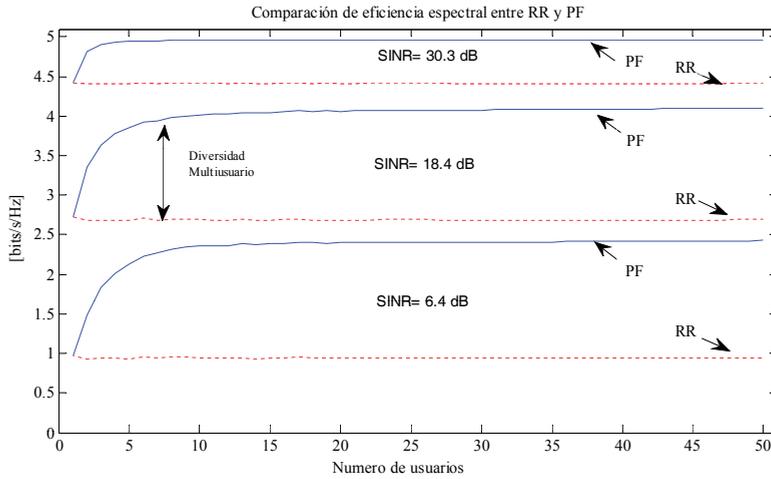


Figura 6-7. Influencia del tipo de estrategia de "scheduling" en la explotación de la diversidad multiusuario

6.4.6 Scheduling semi-persistente

Los mecanismos de *scheduling* que se han descrito hasta el momento permiten que el eNB pueda hacer una planificación de las transmisiones totalmente dinámica y flexible, adaptada a las condiciones cambiantes del canal y el tráfico en la celda. No obstante, esto requiere de una señalización explícita de las asignaciones de *scheduling* que consumen recursos del PDCCH. En el caso de transmisiones de grandes volúmenes de datos, esta señalización puede ser asumible, pero para servicios con pequeños volúmenes de tráfico que llegan de forma regular (por ejemplo VoIP) el *scheduling* dinámico puede ser ineficiente desde el punto de vista de la sobrecarga de señalización. Para ello, el estándar introduce el mecanismo de "*scheduling* semi-persistente" [10].

El *scheduling* semi-persistente permite indicar a un terminal que una asignación de *scheduling* (*downlink scheduling assignment* o *uplink grant*) es válida en lo sucesivo en periodos de, en general, T subtramas, no siendo necesario enviar explícitamente asignaciones de *scheduling* a través del PDCCH a partir de ese momento. Esto es válido mientras no se indique lo contrario por parte del eNB. La decisión de la periodicidad T y la activación o desactivación del *scheduling* semi-persistente reside en las capas RRC.

Adicionalmente, el terminal puede seguir recibiendo asignaciones de *scheduling* dinámicas en cualquier subtrama (en caso de coincidir con una subtrama de *scheduling* semi-persistente, la asignación dinámica tiene preferencia). De este modo, sigue siendo posible variar de forma dinámica los recursos asignados a un terminal.

6.5 Inter-Cell Interference Coordination

La interferencia en los sistemas celulares se puede clasificar entre *intracelular* cuando se

produce entre transmisiones de/para usuarios servidos por una misma celda e *intercelular* cuando la interferencia afecta a transmisiones de/para usuarios servidos por distintas celdas.

El sistema LTE está libre de interferencia intracelular. Como se ha visto en el apartado anterior, la función de *scheduling* asegura que un PRB sólo puede ser asignado a las transmisiones de un único usuario en cada subtrama, evitando colisiones entre las transmisiones de múltiples usuarios al acceder al canal compartido (PDSCH o PUSCH). Sin embargo, en LTE existe, de forma general, interferencia intercelular entre transmisiones de/para usuarios con el mismo PRB asignado en la misma subtrama en distintas celdas (especialmente si son adyacentes). Esta situación puede degradar en exceso las prestaciones del sistema por, entre otras, las siguientes causas:

- Disminuye la SINR por PRB, con lo que se deben seleccionar esquemas de modulación y codificación de mayor protección y de baja eficiencia espectral, reduciéndose así la cantidad de bits que se pueden enviar por PRB.
- Aumenta la potencia de transmisión en UL para asegurar una recepción adecuada, con el consiguiente consumo de potencia en el terminal.
- Se necesita, en general, más retransmisiones HARQ para la entrega favorable de paquetes, incrementándose la carga de señalización.

En sistemas anteriores a LTE como GSM, la interferencia intercelular se mitigaba principalmente en la fase de despliegue y planificación de frecuencias, donde se empleaban técnicas de agrupación de celdas y factores de reuso de frecuencias, que permanecían inalterados durante la operativa del sistema. No obstante, el uso de estas técnicas reduce considerablemente la capacidad potencial en cada celda. Por esta razón, y dado que LTE persigue obtener altas velocidades de pico y eficiencia espectral, LTE se desplegará, presumiblemente, con un reuso universal de frecuencia, donde todos los PRBs estarán disponibles en todos los eNB, maximizando la capacidad potencial por celda. No obstante con este despliegue la interferencia intercelular podría llegar a ser muy elevada, especialmente entre celdas adyacentes y para usuarios alejados del eNB. Ante esta situación, LTE define una función de gestión de recursos radio denominada *InterCell Interference Coordination* (ICIC) cuya misión es gestionar, de forma dinámica, y durante la operativa del sistema, la interferencia intercelular.

La Figura 6-8 muestra el esquema funcional para la aplicación de ICIC en LTE. La funcionalidad ICIC reside en cada eNB pero, como se verá a continuación, para su correcto funcionamiento requiere de la interacción de varios eNBs, tratándose así de una funcionalidad de gestión de recursos radio multicelda. El objetivo ICIC es decidir un conjunto de restricciones y preferencias sobre el uso y la potencia de transmisión por PRB de un determinado eNB para mitigar la interferencia intercelular. Estas restricciones y preferencias se tendrán en cuenta por la función de *scheduling* de paquetes a la hora de planificar las transmisiones de los usuarios en cada eNB. Así, mediante una adecuada selección de las restricciones y preferencias se puede conseguir una cierta coordinación entre los *schedulers* de paquetes en distintos

eNBs. Para ello, las estrategias de ICIC consideran una serie de *indicadores de interferencia* estandarizados que se intercambian los eNBs, así como información (no estandarizada) proveniente del propio eNB (e.g., el nivel de interferencia o decisiones de *scheduling*). Al igual que con otras funciones RRM, el estándar no define las estrategias específicas de decisión acerca de las restricciones/preferencias sobre los PRBs. No obstante, a lo largo del apartado se mencionarán algunas propuestas al respecto. Finalmente, nótese que la ICIC en LTE se hace en el dominio de la frecuencia (mediante restricciones/preferencias sobre los PRBs) en lugar de en el dominio temporal (por ejemplo mediante la alternancia temporal de actividad en las transmisiones en distintos eNBs), ya que este último afectaría a los procesos de HARQ, especialmente en el UL, donde se usa HARQ síncrono (véase el Anexo 5.3 del Capítulo 5).

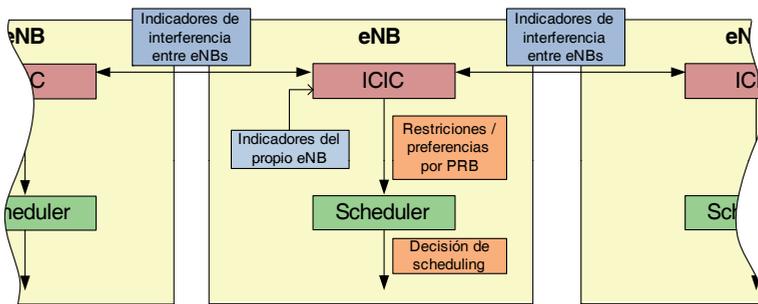


Figura 6-8. Esquema funcional para la aplicación de ICIC en LTE

Los indicadores de interferencia están definidos en [12]. Estos indicadores se intercambian a través de la interfaz X2 entre eNBs (Capítulo 2). En LTE los indicadores de interferencia se clasifican como reactivos y proactivos. Por un lado, los indicadores reactivos se basan en medidas del pasado, de forma que si un eNB detecta una interferencia intercelular muy elevada se lo notifica a las celdas colindantes, a fin de que éstas reduzcan la interferencia generada. Por otro lado, los indicadores proactivos sirven para que un eNB notifique con antelación a los eNBs vecinos el uso que va a hacer de los PRBs en frecuencia, es decir, con qué potencia planea transmitir en cada PRB y dónde pretende planificar las transmisiones de los usuarios en el borde de la celda, que son los que generarán más interferencia. Con esta información, las celdas vecinas pueden predecir la interferencia intercelular que el eNB va a generar a corto plazo.

La Figura 6-9 ilustra el uso de indicadores reactivos y proactivos. Se considera una geometría particular para el enlace ascendente de dos usuarios pertenecientes a las celdas "A" y "C" respectivamente, situados en el borde de sus celdas tal y como muestra la figura. El usuario en la celda "C" hace uso del conjunto de PRBs y_i . Por lo tanto, dada su situación genera una alta interferencia en el receptor del eNB de la celda "A" en dicho conjunto de PRBs y_i . Por otro lado, el *uplink scheduler* del eNB de la celda "A" decide que el usuario de la celda "A" va a transmitir en el conjunto de PRBs x_i , y con una alta potencia debido a su situación lejana,

generando consecuentemente una fuerte interferencia intercelular en la celda "B". En esta situación, los indicadores de interferencia funcionarían como sigue. Por un lado, el indicador reactivo hace que la celda "A" informe a las celdas vecinas ("B" y "C") de que percibe una fuerte interferencia en los PRBs y_i a fin de que éstas intenten reducir la interferencia en estos PRBs. Por otro lado, el indicador proactivo hace que la celda "A" informe también a las celdas vecinas de que en un futuro próximo existirá transmisión con alta potencia en los PRBs x_i . De este modo que las celdas adyacentes podrían evitar usar esos PRBs. Por lo tanto, mediante estos indicadores las estrategias de ICIC pueden decidir una serie de restricciones y preferencias que harán que los eNBs puedan coordinar sus transmisiones para evitar la interferencia intercelular.

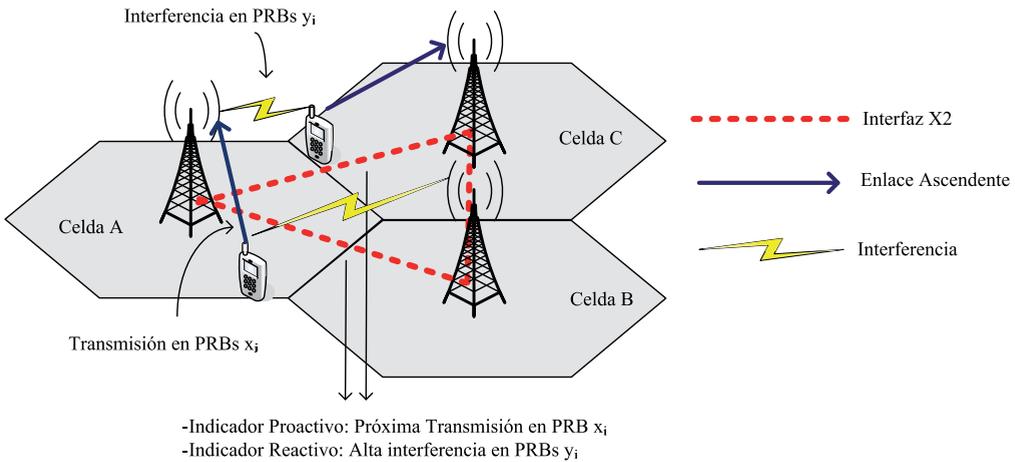


Figura 6-9 Indicadores de interferencia proactivos y reactivos en LTE

6.5.1 Indicadores de interferencia y estrategias ICIC en downlink

El indicador de interferencia estandarizado en DL se denomina *Relative Narrowband Transmit Power* (RNTP) [12]. Se trata de un indicador proactivo que indica, de forma anticipada, la máxima potencia de transmisión en DL por PRB. El RNTP es un vector con tantos bits como PRB disponibles en el sistema que cada eNB envía a sus vecinas a través de la interfaz X2. Si un bit está a 0, las celdas vecinas saben con certeza que el eNB no va a transmitir en ese PRB o lo hará con un nivel de potencia por debajo de un determinado umbral $RNTP_{threshold}$. Esto permitiría a las estrategias de ICIC otorgar cierta preferencia a la planificación de transmisiones en esos PRBs que en principio experimentarían baja interferencia. Por el contrario, si el bit está a 1, los eNBs vecinos no tienen garantías de que la interferencia en esos PRBs sea baja, ya que el eNB de interés podría superar el $RNTP_{threshold}$. Por ejemplo, el eNB pondría este bit a 1, si planea hacer *scheduling* de usuarios situados en el borde de la celda en el PRB correspondiente y a 0 si los usuarios están próximos.

Con el RNTP y el $RNTP_{threshold}$ es posible establecer una estrategia de ICIC que fuera seleccionando de forma dinámica distintos factores de reuso de frecuencias (*Frequency Reuse Factor: FRF*), según las condiciones de interferencia intercelular detectadas en el sistema. Es importante ver que no se trataría de un reuso estático como se hacía en GSM, si no que cada eNB tiene potencialmente todo el espectro disponible en caso de necesitarlo, y que el reuso de frecuencias se obtiene mediante el establecimiento *on-line* de una serie de restricciones sobre el uso de PRBs.

La Figura 6-10 muestra las diversas posibilidades que se han planteado hasta el momento en la línea anterior. Por un lado se ilustra el reuso 1 (FRF1), donde todos los PRB están disponibles en todas las celdas. En el otro extremo está el reuso 3 (FRF3), donde los PRB se dividen en 3 grupos que se asignan a agrupaciones de 3 celdas. Este esquema es muy útil en situaciones donde interese reducir la interferencia intercelular, aunque se reduce drásticamente la capacidad potencial por celda, y en algunos casos se podría no garantizar la calidad de servicio de las conexiones establecidas en términos de velocidad de pico y *throughput* obtenido por los usuarios. Como solución intermedia, se proponen [13][14] un par de estrategias denominadas *partial reuse* (PR) y *soft-frequency reuse* (SR). Tal y como se observa en la Figura 6-10, estas estrategias combinan los esquemas FRF1 y FRF3 en la misma celda, de forma que el esquema FRF1 se usa para los usuarios cercanos al eNB y el FRF3 se deja para los usuarios en el borde de la celda. Nótese que SR además establece niveles de potencia intermedios por PRB, y que todos los PRB están disponibles en cada celda.

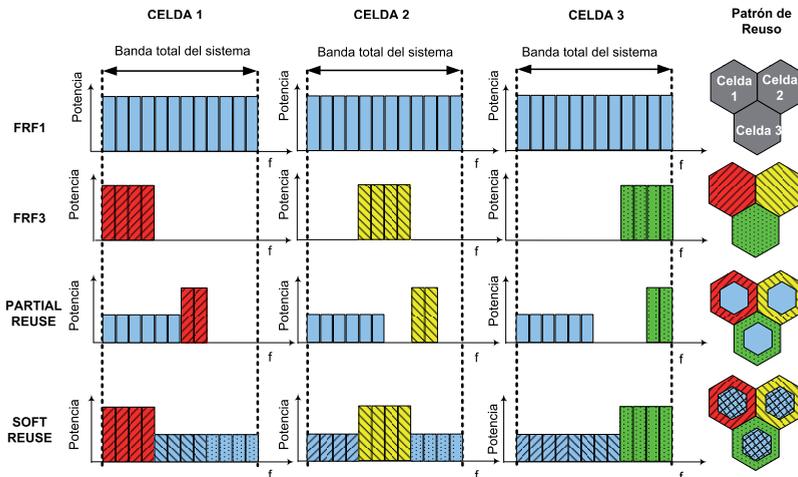


Figura 6-10 Posibles esquemas de reuso de frecuencias en LTE implementados a través de una estrategia ICIC

6.5.2 Indicadores de interferencia y estrategias ICIC en uplink

En el caso del UL, se ha estandarizado un indicador proactivo y otro reactivo. El indicador proactivo se denomina *High Interference Indicator* (HII). Este indicador consiste en un mensaje con un bit por PRB que la celda servidora envía a los eNBs vecinos, indicando (cuando el bit está a 1) la intención de usar esos PRBs para planificar las transmisiones de los usuarios en el borde de la celda (es decir, de aquellos que generan más interferencia en UL). Por otro lado, el indicador reactivo se denomina *Overload Indicator* (OI), y es un mensaje que indica para cada PRB si la interferencia más ruido medida por el eNB es alta, media, o baja [12].

Por su parte, las estrategias ICIC usan esta información para determinar la forma más propicia de planificar las transmisiones de sus usuarios, de forma que se eviten PRBs con una alta interferencia reportados en el HII, y se reduzca la interferencia causada a la celda que inicialmente envió el mensaje OI. La interacción entre HII y OI, y cómo una celda usa esta información para realizar las acciones concretas de ICIC no está estandarizada y posiblemente será dependiente de cada fabricante. Dentro del 3GPP se han hecho varias propuestas que consisten en provocar que las estrategias de *scheduling* planifiquen a los usuarios en zonas del espectro complementarias a aquellas en la que se ha reportado una alta interferencia con el HII [15][16]. Por otro lado, también se puede usar el mensaje OI para ajustar de forma dinámica el control de potencia en lazo abierto, de forma que se mantenga un determinado nivel de interferencia más ruido en la celda (ajuste del parámetro P_0 tal y como se define en el apartado 1.8).

6.5.3 Otras técnicas de mitigación de interferencia intercelular

Al margen de ICIC, existen otras propuestas de estrategias que podrían reducir la interferencia intercelular en LTE [17]. Algunas de estas estrategias están basadas en el uso del control de potencia [18], donde se controla la potencia transmitida en UL para compensar sólo una fracción de las pérdidas de propagación radio o *pathloss* (esto es, las técnicas de *partial pathloss compensation* que se describirán en el subapartado 1.8.2). Otras propuestas se basan en técnicas de acceso por división espacial [19] y *beamforming* [20], donde se crean haces selectivos en los diagramas de radiación de las antenas para focalizar la radiación y así reducir la interferencia intercelular. También existen otras propuestas basadas en técnicas de procesado y estimación de la señal cuyo objetivo es cancelar la interferencia [21]. La idea es detectar y regenerar la interferencia para posteriormente sustraerla de la señal recibida. Por último, recientemente se han propuesto técnicas de codificación que permiten mitigar el efecto de la interferencia. Algunas referencias en este campo son [22][23].

6.6 Control de la movilidad

La movilidad es una de las funcionalidades clave en los sistemas de comunicaciones móviles celulares, ya que permite a los usuarios acceder y recibir servicio desde cualquier ubicación geográfica donde el sistema disponga de cobertura. Para ello, el sistema proporciona una serie de procedimientos de movilidad que permite a los usuarios moverse con continui-

dad del servicio por el área geográfica cubierta. Como se describió en el Capítulo 3, en LTE estos procedimientos se dividen básicamente en dos categorías: los de reelección de celda y actualización de la localización cuando el terminal se encuentra en estado RRC-IDLE (i.e., sin conexiones establecidas) y los de traspaso o *handover* cuando el terminal se encuentra en estado *RRC-CONNECTED* (i.e., con conexiones establecidas con el eNB). No obstante, al margen de los procedimientos proporcionados por el sistema, se hace necesaria la implementación de una serie de estrategias RRM de control de la movilidad que permitan la adecuada monitorización del área de cobertura en la que están registrados los terminales³, así como la selección del instante de ejecución de dichos procedimientos en caso de que el área de cobertura de la que se obtiene servicio deje de ser la más favorable para comunicarse de manera fiable con el sistema.

En este apartado se describirán dichas estrategias RRM. En el subapartado 6.6.1 se hace hincapié en las tareas que siguen estas estrategias así como en las medidas utilizadas en LTE. Posteriormente, en el subapartado 6.6.2 se presentan estrategias de reelección de celda, mientras que en el subapartado 6.6.3 se describen estrategias de *handover*. Como ya se ha comentado, la descripción de los procedimientos de movilidad asociados y estandarizados en LTE se puede encontrar en el Capítulo 3.

Antes de entrar en la descripción de las estrategias, conviene introducir los escenarios movilidad contemplados en LTE así como las situaciones en las que se contempla que debe actuar la gestión de la movilidad [2]. En lo que respecta a los escenarios, se trata de tres escenarios denominados *intra-frequency*, *inter-frequency*, e *inter-RAT (Radio Access Technology)* que se describen a continuación.

El escenario *intra-frequency* hace referencia a eventos de movilidad que implican un cambio de celda entre celdas que están desplegadas en la misma banda de frecuencias (es decir, que usan la misma portadora LTE). Suele ser el escenario más habitual.

El escenario *inter-frequency* hace referencia a despliegues del operador donde existen múltiples portadoras o bandas de frecuencias LTE trabajando en paralelo. Por ejemplo, en un despliegue jerárquico con macro-, micro- y femto- celdas donde cada tipo de celda usa una banda de frecuencias distinta. Estas bandas no tienen por qué ser necesariamente contiguas e incluso podrían ser antiguas bandas de UTRAN o GERAN que pasan a ser de LTE a medida que LTE entra en su fase de maduración.

Finalmente, el escenario *inter-RAT* considera aspectos similares a los del *inter-frequency* pero esta vez en un despliegue en el que existen múltiples tecnologías de acceso radio (por ejemplo GERAN, UTRAN y LTE).

Entre los motivos (*drivers*) que pueden provocar un evento de movilidad en estos escenarios hay varios, tal y como recoge la Tabla 6.1 [2]. De ellos, el motivo más relevante suele ser el de mantener la mejor condición de propagación para el terminal en cualquiera de los

³ Entendiendo aquí el área de cobertura a nivel de un conjunto de celdas (área de localización) cuando el terminal se encuentra en estado RRC-IDLE o a nivel de una celda individual cuando el terminal se encuentra en estado RRC-CONNECTED.

escenarios presentados, a fin de preservar una conectividad fiable con el sistema. Por otro lado, para los escenarios *inter-frequency* e *inter-RAT* existen otros motivos adicionales que pueden desencadenar la ejecución de los procedimientos de movilidad como son, el balanceo de carga, las capacidades del terminal, el tipo de servicio, el despliegue del sistema, etc. Finalmente, existen algunos condicionantes que se debe tener en cuenta a la hora de diseñar las estrategias de control de la movilidad y que también están recogidos en la Tabla 6.1.

Tabla 6.1 Drivers y limitaciones de la movilidad en LTE [2]

	#	Drivers/Limitaciones	Intra-frequency	Inter-frequency	Inter-RAT
Drivers	1	Mejor condición de propagación	X	X	X
	2	Camp load balancing		X	X
	3	Traffic load balancing		X	X
	4	Capacidades del terminal		X	X
	5	Despliegues Jerárquicos		X	X
	6	Network sharing		X	X
	7	Redes privadas/caseras		X	X
	8	Tipo de subscripción		X	X
	9	Tipo de servicio		X	X
	10	MBMS		X	X
Limitaciones	11	Ahorro de batería en el terminal	X	X	X
	12	Señalización de red y carga de procesado	X	X	X
	13	Pérdida de datos e interrupción de la conexión	X	X	X
	14	Complejidad en la gestión y mantenimiento	X	X	X

En lo sucesivo nos centraremos en estrategias pensadas para mantener la mejor condición de propagación, mientras que dentro del apartado 6.7, dedicado a la descripción de la funcionalidad RRM de balanceo de carga, se describirá en detalle algunas estrategias de control de la movilidad motivadas por aspectos de balanceo de carga como *camp load balancing* o *traffic load balancing*.

6.6.1 Etapas para la ejecución de estrategias RRM de movilidad.

La Figura 6-11 ilustra las actividades típicas que debe ejecutar una estrategia de control de la movilidad en una red de comunicaciones móviles celulares, donde las etapas se ejecutarán en el terminal o en el eNB dependiendo de si el terminal se encuentra en estado *RRC-IDLE* o *RRC-CONNECTED*, como se detallará en los subapartados sucesivos.

Se parte de un estado en que el terminal monitoriza y evalúa periódicamente la calidad de la señal recibida desde la celda servidora. Si la calidad está por encima de un determinado

umbral, entonces el terminal continúa en este estado. Si por el contrario la calidad de la celda servidora es insuficiente, se interpreta que puede ser necesaria la ejecución de un procedimiento de movilidad y se ejecutan las siguientes etapas.

En primer lugar, el terminal hace una búsqueda de celdas distintas de la servidora en sus proximidades y elabora una lista. Esta lista podría contener celdas LTE (*intra-frequency* o *inter-frequency*) o celdas de otros sistemas (*inter-RAT*). La búsqueda se irá repitiendo periódicamente hasta que se ejecute un procedimiento de movilidad o mejoren las condiciones en la celda servidora. En LTE, este procedimiento de búsqueda es similar al de sincronización y adquisición inicial de parámetros explicado en el apartado 5.8 del Capítulo 5. Básicamente, consiste en adquirir sincronización de frecuencia y símbolo a una celda, alinearse temporalmente con ella y determinar la identidad física de la celda (PCI). Con esta identidad física, el terminal puede identificar unívocamente una señal de referencia específica (*Reference Signal*) de la celda y comenzar la estimación de la intensidad y calidad de la señal recibida. El terminal puede seguir procedimientos similares para otras celdas de otras RATs como UTRAN o GERAN. No obstante, para facilitar este proceso y así reducir el consumo del terminal, la red puede asistir al terminal mediante el envío de listas (*Neighbour Cell List: NCL*) en las que centrarse para realizar la búsqueda. En el caso de UTRAN la NCL incluye hasta 32 celdas con sus respectivos códigos de *scrambling*⁴ y se envía en el "*System Information Block*" número 6 (SIB6) descrito en el Anexo 5.3 del Capítulo 5. Por otro lado, en el caso de GERAN, la NCL (enviada por el SIB7) incluye hasta 32 frecuencias portadoras de las celdas vecinas y opcionalmente sus respectivos indicadores de celda⁵.

Una vez que se tiene una lista de celdas candidatas se pasa a tomar medidas de la intensidad y calidad de la señal recibida, que permitirán a la siguiente etapa evaluar si es o no necesaria la ejecución de un procedimiento de movilidad. En LTE básicamente se consideran dos medidas que están definidas en [24]⁶:

- **Reference Signal Received Power (RSRP):** Esta medida proporciona una estimación de la fuerza de la señal recibida a nivel de celda, y, como veremos, es útil para la toma de decisiones acerca de la ejecución de procedimientos de tanto reelección de celda como de *handover*. El RSRP se mide como el promedio lineal sobre las potencias recibidas (en vatios) de los *resource elements*⁷ que transportan la *Reference Signal* de una celda determinada en un ancho de banda especificado. Para poder tomar esta medida, el terminal debe estar sincronizado con la celda objetivo (es decir, debe adquirir los canales P-SCH y S-SCH vistos en el Capítulo 5).
- **Reference Signal Received Quality (RSRQ):** Esta medida trata de proporcionar una estimación de la calidad recibida desde una celda en particular. También se puede usar

⁴ Los códigos de scrambling son secuencias pseudoortogonales que se usan en W-CDMA para diferenciar la información enviada por distintas celdas.

⁵ En concreto se incluye el Absolute Radio Frequency Channel Number (ARFCN) como indicador de la frecuencia portadora y el Base Station Identity Code (BSIC) como identificador de celda.

⁶ Para UTRAN y GERAN se considera las medidas estandarizadas para esos sistemas.

⁷ Un resource element es una subportadora OFDM en un determinado símbolo OFDM dentro de un time-slot (véase el Capítulo 5).

como entrada para los algoritmos de decisión de reelección de celda y *handover*. Se define como

$$RSRQ = \frac{N * RSRP}{RSSI}, \quad (6.4)$$

donde RSSI (*Received Signal Strength Indicator*) es la potencia de portadora LTE observada por el terminal incluyendo la potencia recibida de la celda servidora y celdas no servidoras en la misma portadora LTE, la interferencia de canal adyacente y el ruido térmico. Por otra parte, N es el número de PRBs donde se mide el tanto el RSRP y el RSSI. Por lo tanto, el RSRQ es una estimación de la relación señal a ruido más interferencia observada por el terminal para una determinada celda objetivo.

Posteriormente, el algoritmo de decisión se encargará de decidir si es necesario ejecutar un procedimiento de movilidad en base a unos determinados criterios definidos a partir de las medidas tomadas.

Por último, conviene subrayar que el resultado de cada una de las etapas descritas dependerá de la configuración de determinados parámetros definidos por el operador (e.g., umbrales, márgenes de histéresis, tiempos de promediado, etc.). Obviamente la configuración de estos parámetros también forma parte de la gestión de recursos radio de los eventos de movilidad.

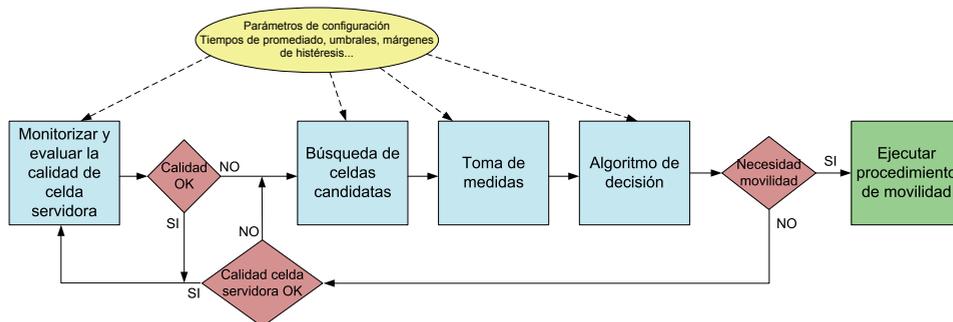


Figura 6-11 Etapas para la ejecución de estrategias RRM de movilidad

6.6.2 Estrategias de reelección de celda

El objetivo de las estrategias de reelección de celda es asegurar que el terminal esté registrado (acampado) en la celda más favorable cuando se encuentra en estado RRC-IDLE. En este estado, el terminal se limita a comprobar periódicamente si tiene mensajes de búsqueda (*paging*) para él y a monitorizar el estado de la señal recibida de la celda que considera como servidora, entendiendo como tal a aquella a la que el terminal se encuentra 'enganchado' y

decodifica su información del canal de *broadcast* (BCH)⁸. Esta información incluye información específica de la celda como configuración *paging* y acceso aleatorio, y configuración para realizar medidas y para los algoritmos de reelección de celda.

En LTE, las estrategias de reelección de celda se ejecutan de manera autónoma en el terminal, de forma que éste se encarga de monitorizar la celda servidora, buscar otras celdas, tomar medidas y decidir la nueva celda en caso necesario. De esta forma se consigue reducir la señalización en la interfaz radio, evitando continuas comunicaciones entre el terminal y el eNB. De hecho, el terminal podrá ir cambiando la celda servidora (mediante reelección de celda) pero no se comunicará con la red mientras no sea necesario el cambio de un área de seguimiento (*Tracking Area*). Por otro lado, esta reducción en señalización aporta una reducción en el consumo por parte del terminal, alargando la vida de la batería.

Una vez que el terminal se enciende y se registra en el sistema, necesita acampar por primera vez en una celda elegida a través de las llamadas estrategias de *cell selection* [25]. Para ello, el terminal selecciona la celda con mayor señal, medida a través de la denominada medida S , definida como [25]:

$$S = Q_{\text{rxlevelmeas}} - (Q_{\text{rxlevmin}} + Q_{\text{rxlevelminoffset}}) - P_{\text{compensation}}, \quad (6.5)$$

donde $Q_{\text{rxlevelmeas}}$ es el nivel de señal recibido (RSRP en el caso de LTE), Q_{rxlevmin} es un nivel mínimo de señal recibido en la celda en dBm, y $Q_{\text{rxlevelminoffset}}$ y $P_{\text{compensation}}$ son parámetros configurables por el operador.

A partir de ahí, el terminal está continuamente monitorizando la calidad del canal con la celda servidora a través de la medida S . Siempre que esta medida sea lo suficientemente alta, entonces el terminal no realizará ninguna medida adicional y permanecerá acampado en la celda servidora. Si por el contrario, la medida S cae por debajo de unos determinados umbrales, entonces se iniciará el proceso de búsqueda y selección de una nueva celda tal y como se explica a continuación.

Por un lado, si la medida S para la celda servidora ($S_{\text{servingcell}}$) es menor o igual que el umbral $S_{\text{intrasearch}}$, entonces se inicia la búsqueda de celdas *intra-frequency*. Para ello, se elabora un *ranking* entre la celda servidora y otras celdas que se basa en las cantidades R_s para la primera y R_n para el resto, siendo

$$\begin{aligned} R_s &= Q_{\text{meas},s} + Q_{\text{hyst}}; \\ R_n &= Q_{\text{meas},n} + Q_{\text{offset}}. \end{aligned} \quad (6.6)$$

Q_{meas} es el RSRP de la celda en cuestión, Q_{hyst} es un margen de histéresis en potencia y Q_{offset} es un *offset* específico para controlar el comportamiento de la estrategia de reelección de celda con una celda específica (e.g., para motivar la selección de una determinada

⁸ No obstante, como se vio en el Capítulo 3, en estado RRC-IDLE la red sólo conoce la localización de un terminal a nivel de un conjunto de celdas denominado “Tracking Area”.

celda sobre otras). Los dos criterios se promedian durante $T_{reselection}$ muestras y si algún R_n supera a R_s durante $T_{reselection}$ muestras, entonces se selecciona la nueva celda. La Figura 1-12 ilustra este criterio.

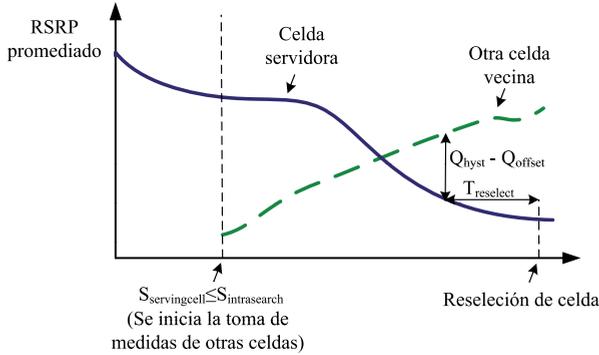


Figura 6-12 Proceso del criterio de re-selección de celda intra-frecuencia en LTE

Por otro lado, si $S_{servingcell}$ es menor o igual que $S_{intersearch}$, se inicia la búsqueda de celdas *inter-frequency* e *inter-RAT*. En este caso, LTE introduce (como novedad con respecto a sistemas 3GPP anteriores) una serie de prioridades absolutas para las frecuencias de portadora de LTE o de otras RATs. El listado de prioridades lo obtiene el terminal a través de la información de sistema enviada en el canal de *broadcast* o por señalización RRC cuando el terminal pasa a estado *RRC-IDLE* desde *RRC-CONNECTED*, pudiendo ser en este caso, una lista de prioridades específica para cada terminal. Este nuevo mecanismo de prioridades permite al operador establecer ciertas preferencias sobre frecuencias y RATs (en definitiva sobre las celdas) en las que quiere que acampen los terminales. Básicamente, el objetivo es seleccionar la celda de la frecuencia y RAT de mayor prioridad que cumpla con unos mínimos de calidad e intensidad de la señal recibida. Para ello, la red establece un umbral, $Thresh_{x, high}$, que debe superar la medida S de cualquier celda de una frecuencia o RAT x de mayor prioridad durante $T_{reselection, RAT}$ muestras antes de seleccionarla. $T_{reselection, RAT}$ es similar al parámetro $T_{reselection}$ pero específico para cada RAT. En caso de que no haya disponibles celdas de una RAT de mayor prioridad y sólo haya celdas LTE de frecuencias de igual prioridad, se ejecuta el mismo criterio que para re-selección de celda *intra-frequency* descrito anteriormente. Por último, si sólo hay celdas de RATs y frecuencias de menor prioridad, se ofrece continuidad del servicio permitiendo la selección de alguna de esas celdas si se cumple que la medida S está por encima de un umbral $Thresh_{x, low}$ durante, también durante $T_{reselection, RAT}$ muestras.

6.6.3 Estrategias de handover

En el caso de que el terminal se encuentre en modo *RRC-CONNECTED*, la red se encarga de decidir cuando un terminal tiene que cambiar de celda para mantener la calidad del enlace radio (es decir, se ejecuta un procedimiento de *handover* HO).

Para poder decidir cuándo hay que hacer un HO para un terminal, la red necesita medidas de ese terminal acerca de la señal recibida desde su celda servidora y desde otras celdas en las proximidades. Para ello, la red configura al terminal (a través de mensajes de la capa RRC) para que éste monitorice y calcule determinadas medidas de celdas vecinas y envíe los resultados en unos determinados mensajes, conocidos como *measurement reports*.

El procedimiento para monitorización y cálculo de medidas en LTE están definidos en [26]. La red configura en el terminal un determinado patrón de instantes de monitorización de la señal proveniente de otras celdas tanto *intra-frequency* como *inter-frequency* e *inter-RAT*. Este patrón está definido por, (1) una determinada duración del tiempo de monitorización (*gap*), (2) la periodicidad con que se debe repetir cada *gap*, y (3) el tiempo mínimo que se debe tomar una determinada medida para promediarla adecuadamente. Es importante destacar que durante el tiempo en que el terminal está monitorizando la señal de otras celdas se interrumpe cualquier actividad de transmisión y recepción con la celda servidora. Por lo tanto, existe un compromiso entre el desempeño del procedimiento de toma de medidas y la capacidad del terminal para transmitir y recibir datos.

La monitorización de celdas se ilustra en la Figura 1-13. La duración del *gap* en LTE es de 6 ms, y la periodicidad un múltiplo de la trama LTE, es decir 10 ms. En la monitorización *intra-frequency* e *inter-frequency*, el terminal mide el RSRP y el RSRQ de otras celdas en cada *gap*. Para ello, necesita decodificar el *Primary Synchronization Channel* (P-SCH) y el *Secondary Synchronization Channel* (S-SCH) (véase Capítulo 5) de cada celda con la consecuente sincronización. Las medidas en el caso de *intra-frequency* se promedian durante 200 ms. En el caso de *inter-frequency* el tiempo de promediado es mayor, ya que de los 6 ms de duración de cada *gap* se pierde algún tiempo (aproximadamente 1 ms) en re-sintonizar el receptor a la correspondiente frecuencia, quedando menos tiempo para la toma de medidas. Por lo tanto, se debe promediar durante más *gaps*. Finalmente, en la monitorización inter-RAT, se usa la misma duración de *gap*, pero lógicamente se monitorizan otros canales y se toman otras medidas dependientes de la RAT en cuestión [26].

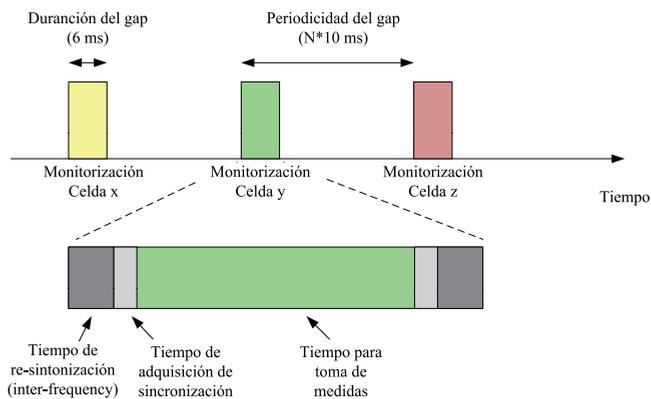


Figura 6-13. Monitorización de otras celdas para la toma de medidas en LTE

En lo que respecta a la actividad de reporte de medidas (*measurement reporting*), está definida en [27]. En el caso de LTE, se reporta el valor del RSRP o RSRQ, mientras los valores reportados para otras RATs dependen de cada RAT (e.g. la relación energía de chip a densidad de ruido térmico del canal piloto, E_c/N_0 , en caso de UTRAN). Por otro lado, existen dos formas de hacer el envío de *measurement reports*: Periódica y a demanda (*event-triggered*). Ésta última se emplea para reducir la señalización entre el terminal y el eNB, haciéndose el envío sólo cuando se cumplen unas determinadas condiciones que podrían estar alineadas con las condiciones que usa la red para determinar que se hace necesario un HO como se verá a continuación.

Los algoritmos de decisión de ejecución de HO en el eNB pueden entrañar varios criterios. El más habitual es el de preservar la calidad de la señal recibida por el terminal, de forma que esté conectado a la celda de mayor calidad. La Figura 6-14 ilustra este procedimiento, donde se emplea el RSRP promediado y reportado por el terminal. Si el RSRP promediado para alguna de las celdas adyacentes es superior al RSRP promediado de la celda servidora más un cierto margen de histéresis (HOM) entonces se inicia un contador denominado "*Time-To-Trigger*" (TTT). Finalmente si la condición se mantiene hasta que expira el TTT, entonces la red determina que es necesario un HO para el terminal. Lógicamente, la configuración de estos parámetros (HOM, TTT) y otros específicos de la estrategia de HO determinan en gran medida el resultado de la gestión de la movilidad en LTE y en líneas más concretas la tasa de HO en la red y la probabilidad de pérdida o caída de la conexión durante el HO. Por ello, se han llevado a cabo diversos estudios sobre la optimización y control del HO en LTE. Por ejemplo, [28] presenta un modelo empírico para la predicción y decisión del HO, mientras que [29] recomienda un rango de HOM en dB considerando en número medio de HO para distintas velocidades del terminal. Por otra parte, las referencias [30] y [31] estudian la configuración del tiempo de promediado y el uso de técnicas de coordinación de la interferencia para optimizar el HO.

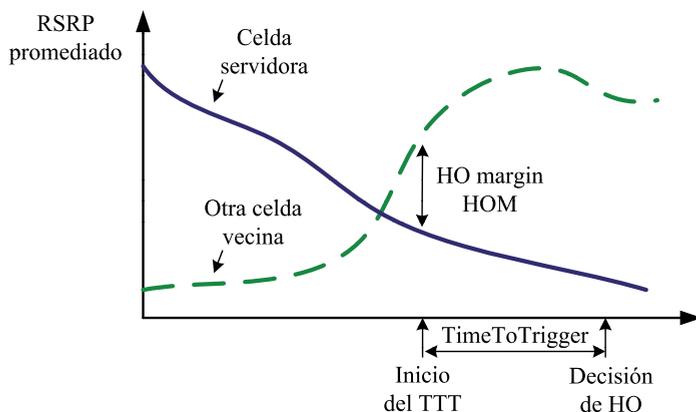


Figura 6-14 Parámetros utilizados para la decisión de HO en LTE

Finalmente, conviene destacar que es posible que el operador esté interesado en realizar HO por otros motivos distintos del mantenimiento de la calidad de la señal recibida. Por

ejemplo, el operador podría estar interesado en derivar un determinado tipo de servicio a una determinada RAT, en la línea del *driver* número 9 en la Tabla 6.1. También, el operador podría querer que las transmisiones de los terminales se distribuyeran homogéneamente entre distintas celdas o bandas de frecuencias, dando pie a estrategias de balanceo de carga que se verán en el siguiente apartado.

6.7 Balanceo de carga

Durante la planificación y despliegue de una red celular se asignan recursos a cada estación base de forma acorde a la demanda de tráfico esperada. Esto lleva a un reparto similar de recursos radio por celda en áreas de características de tráfico similares. No obstante, debido a las condiciones cambiantes de un sistema celular, es posible que se llegue a situaciones de descompensación de la carga, donde unas estaciones base estén sobrecargadas (es decir, cursando más tráfico del que pueden soportar) y otras en la misma zona estén trabajando por debajo de su capacidad.

El Balanceo de carga, *Load Balancing* (LB) es un mecanismo que se encarga de gestionar distribuciones no uniformes de usuarios (y por lo tanto de solicitudes del servicio) entre múltiples celdas [2]. En LTE, LB se contempla tanto para usuarios que se encuentran en estado RRC-IDLE, como para aquellos que tienen conexiones activas en estado *RRC-CONNECTED*. El primero atiende al motivo de movilidad *camp load balancing* en la Tabla 6.1, mientras que el segundo hace referencia al motivo de movilidad *traffic load balancing* también incluido en esa tabla, ambos explicados en lo sucesivo.

- El procedimiento de *Camp load balancing* busca distribuir uniformemente a los usuarios en estado no activo (RRC-IDLE) entre las bandas/portadoras/RATs disponibles. De esta forma se consigue que, tras la activación, la carga de tráfico de las bandas/portadoras/RATs esté balanceada. También se iguala el consumo de recursos radio por parte de los terminales a la hora de responder a llamadas de *paging* o actualizar el área de seguimiento. Por último, estas acciones de balanceo pueden verse motivadas por tener a determinados usuarios, e.g., aquellos con suscripciones de alta velocidad de datos, ya sincronizados con la RAN por la que el operador tiene interés en cursar sus servicios, evitando así el tener que hacer un HO cuando entren en estado *RRC-CONNECTED*.

El mecanismo para conseguir el balanceo de carga se basa en el envío de prioridades por banda de frecuencias y RAT visto en el apartado 6.6.2, de forma que el operador orienta a los terminales hacia una determinada banda de frecuencias y RAT (nótese que, tal y como se vio en aquel apartado, la lista de prioridades podría ser específica para cada terminal). Por otro lado, también es posible balancear la carga controlando el parámetro Q_{offset} en las expresiones de *cell reselection intra-frequency* dadas en (6.6) y que es específico para cada celda. Por último, el balanceo de carga puede atender a decisiones de la red troncal. Concretamente, se define un parámetro deno-

minado "RAT/Frequency Selection Priority" (RFSP) que la MME proporciona al eNB a través de la interfaz S1 y que determina las preferencias desde el punto de vista de la red troncal a la hora de establecer prioridades de reelección de celda [6].

- El procedimiento de Traffic load balancing balancea los usuarios que se encuentran en estado activo (RRC-CONNECTED), lo que supone la ejecución de un HO, que normalmente será *inter-frequency* o *inter-RAT*. Para interferir poco en la actividad del usuario, LB se debe tratar de aplicar durante periodos de inactividad del usuario (e.g., durante la recepción discontinua). La Figura 6-15 ilustra el concepto de Traffic load balancing, donde se observa que se produce una distribución de las conexiones entre dos eNBs.

En LTE este procedimiento es esencial debido a la naturaleza compartida del canal, ya que un incremento en el número de usuarios activos supone un descenso del *throughput* por usuario, y por lo tanto de la percepción del servicio del usuario. Algunos estudios [32] revelan que, mediante un ajuste apropiado de los parámetros de HO orientados a conseguir balanceo de la carga, se puede conseguir hasta un 30% de mejora en el *throughput*.

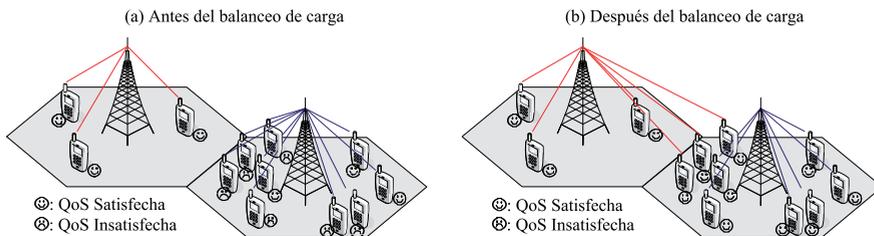


Figura 6-15 Ilustración del mecanismo de "traffic load balancing"

6.8 Control de potencia

El control de potencia es el conjunto de herramientas por las que se controla la potencia de transmisión de los canales físicos para asegurar que se reciben correctamente en cualquier punto de la celda a la vez que permite alargar la vida de la batería de los terminales.

En el caso del DL, el control de potencia estaría totalmente controlado por el eNB, ya que conoce la potencia disponible en todo momento para la transmisión. No obstante, lo más común es que no se implemente control de potencia, ya que, debido a que el consumo de potencia no es crítico para el eNB, no es un problema distribuir toda la potencia disponible en el eNB uniformemente entre los PRBs. Existen algunas propuestas para asignación óptima de potencia denominadas *water-filling* que consisten en asignar dinámicamente más potencia a aquellas subportadoras OFDM que tienen las mejores condiciones de desvanecimiento e interferencias pero su ganancia es marginal [33].

A diferencia del DL, en el UL no es recomendable hacer que los terminales transmitan con máxima potencia, ya que esto reduciría drásticamente la duración de la batería e incrementaría la interferencia intercelular. Por ello, el control de potencia en el UL se hace indispensable y es el que se abordará en el resto del apartado.

El control de potencia en el UL en LTE es centralizado y está controlado por el eNB. Como veremos, éste necesita enviar comandos de ajuste de potencia de transmisión al terminal, lo que exige que el mecanismo de control de potencia para el UL tenga que estar estandarizado en LTE. Básicamente, el control de potencia en UL es una combinación de un mecanismo de control en lazo abierto y en lazo cerrado. En lazo abierto (*open-loop*), la potencia de transmisión del terminal se ajusta en función de una estimación de las pérdidas de propagación (*pathloss*) en DL, de forma que se compensen dichas pérdidas para la transmisión en UL. En lazo cerrado (*close-loop*) la red envía al terminal comandos específicos de control de potencia para ajustar y controlar la potencia de transmisión del terminal. Estos comandos se envían al terminal cada TTI (*Transmission Time Interval*) del PDCCH.

En LTE se estandarizan los mecanismos de control de potencia en el UL para los siguientes canales físicos y señales de referencia:

- *Physical Uplink Control Channel (PUCCH)*.
- *Physical Uplink Shared Channel (PUSCH)*.
- *Sounding Reference Signals (SRS)*.

A continuación se detalla el procedimiento para los dos primeros (PUCCH y PUSCH), mientras que el procedimiento para el tercero (muy similar a los anteriores) se puede encontrar en [8].

6.8.1 Control de potencia para el PUCCH

El control de potencia para el canal PUCCH se establece según la siguiente fórmula para cada subtrama:

$$P_{T_PUCCH} = \min \left\{ P_{\max}, PL_{DL} + P_0 + \Delta_{\text{Format_PUCCH}} + \delta \right\}. \quad (6.7)$$

Esta expresión establece en dBm la potencia transmitida por el terminal en una determinada subtrama del P_{T_PUCCH} . P_{\max} es la potencia máxima del terminal (por ejemplo 23 dBm). La primera parte de la expresión ($\min \{ P_{\max}, \dots \}$) establece que, en cualquier caso, la potencia transmitida es inferior a la potencia máxima disponible en el terminal, aunque es recomendable no acercarse a ella para evitar que en determinados momentos la potencia transmitida sea insuficiente para decodificar satisfactoriamente el canal PUCCH en recepción.

Por otro lado PL_{DL} es *pathloss* estimado en DL por los terminales y medido en dB a partir de medidas de la potencia recibida de señales de referencia conocidas por el terminal y

detalladas en el Capítulo 5. P_0 es un parámetro específico de la celda disponible con la información de *broadcast* y establece la potencia (en dBm) que se desea recibir en la estación base. Esta potencia recibida dependerá de la SINR objetivo deseada y por lo tanto debe variar en función del nivel de interferencia en la celda. Adicionalmente, $\Delta_{\text{Format_PUCCH}}$ establece un margen de potencia en la potencia transmitida que depende del formato de modulación y codificación empleado para el PUCCH.

Los términos hasta ahora descritos permitirían establecer un control de potencia en lazo abierto. No obstante, es posible implementar un control de potencia en lazo cerrado que serviría para compensar los desvanecimientos de canal multicamino en UL que no están incluidos en la estimación del *pathloss* en DL, o las variaciones de interferencia que no refleja el parámetro P_0 . El control de potencia en lazo cerrado se implementa mediante el envío de comandos específicos de control de potencia al terminal para ajustar el término δ en (6.7). Cada comando de control de potencia aumenta o disminuye el valor de δ en una determinada cantidad en dB predeterminada según el comando recibido. El terminal recibe estos comandos cada vez que se le asignan recursos explícitamente en DL (en los *downlink scheduling assignments* vistos en el apartado 1.4) o periódicamente a través de un formato especial del PDCCH, que proporciona comandos para varios terminales. En el primer caso, los comandos de control de potencia establecen variaciones de -1 dB, 0 dB, +1 dB o +3 dB. En el segundo caso, las variaciones son sólo de -1 dB y +1 dB.

6.8.2 Control de potencia para el PUSCH

El control de potencia para el PUSCH sigue la siguiente expresión para cada subtrama del PUSCH:

$$P_{T_PUSCH} = \min \left\{ P_{\max}, PL_{DL} + P_0 + \Delta_{\text{Format_PUCCH}} + \delta \right\}. \quad (6.8)$$

donde P_{T_PUSCH} es la potencia transmitida en UL para el canal PUSCH. Esta expresión es muy similar a la del PUCCH salvo por los términos αPL_{DL} y $10 \log_{10}(M)$. El resto de parámetros tiene una interpretación análoga a los de la expresión para el PUCCH, donde ahora el parámetro δ se actualiza a través de los *uplink scheduling grants* (apartado 6.4) para UL en lugar de en los *downlink scheduling assignments*, y con la misma granularidad en dB.

El parámetro δ permite establecer dos posibles configuraciones de compensación del *pathloss*. Con $\alpha = 1$ se tiene la compensación total (*full compensation*), muy similar a la del PUCCH, que asigna un valor de potencia que asegura una determinada SINR para un determinado formato de modulación y codificación elegido para el PUSCH. Con $\alpha < 1$ se tiene compensación parcial (*partial compensation*), donde un incremento de *pathloss* no tiene un incremento en la misma proporción de la potencia transmitida. Esto sirve para reducir la potencia transmitida de los usuarios en el borde de la célula, al mismo tiempo que se reduce su

tasa binaria. Con esto se consigue disminuir la interferencia que introducen estos terminales, constituyendo un mecanismo de control de la interferencia intercelular, tal y como se apuntó en el subapartado 6.5.3.

Por otro lado, el término $10 \log_{10}(M)$ establece que el control de potencia se realiza por PRB donde M es el número de PRBs asignados al terminal para la transmisión en PUSCH en una determinada subtrama. Naturalmente la potencia transmitida aumenta linealmente con el número de PRBs al hacerlo el número de subportadoras utilizadas. Para asistir al *uplink scheduler* en el eNB en la elección del valor M y el formato de modulación y codificación, $\Delta_{\text{Format_PUSCH}}$, el terminal reporta los llamados *power headrooms*. Estos mensajes establecen la diferencia entre la máxima potencia del terminal y la potencia nominal que usaría el terminal siguiendo la ecuación de control de potencia si no estuviese limitado en potencia. Esto es:

$$\text{PowerHeadroom} = P_{\text{max}} - (\alpha PL_{DL} + P_0 + 10 \log_{10}(M) + \Delta_{\text{Format_PUSCH}} + \delta). \quad (6.9)$$

Asumiendo que la red conoce razonablemente PL_{DL} a través de estimaciones del terminal que se envían a través del PUCCH, y conocidos α , P_0 y δ , la red puede establecer el número de PRBs y el formato de modulación y codificación (es decir el *Transport Format*) que se le asigna al terminal en una determinada subtrama, de forma que el terminal tenga potencia suficiente para soportarlos. Nótese que el *power headroom* no es una potencia. De hecho se mide en dB y en valores en el rango [-23,40] dB en saltos de 1 dB. Por lo tanto el *power headroom* podría ser negativo, indicando que el terminal estaría limitado en potencia ya que el segundo término de (6.9) representa la potencia nominal de transmisión que necesitaría el terminal si no estuviese limitado en potencia, mientras que en la ecuación (6.8) sí incluye la limitación en potencia. En ese caso de limitación de potencia en el terminal, el *uplink scheduler* debería asignar un número pequeño de PRBs al terminal para incrementar la densidad espectral de potencia y un TF compatible con una modulación de bajo orden (QPSK), que requiera una SINR baja para su correcta recepción.

6.9 Recepción discontinua

La recepción discontinua (*discontinuous reception: DRX*) es un mecanismo por el cual el terminal desactiva la circuitería de recepción durante un determinado lapso de tiempo con el fin de ahorrar energía y alargar la vida de la batería. Desde el punto de vista del retardo o latencia, sería muy beneficioso tener continuamente encendido el receptor del terminal, ya que se podría monitorizar en cada subtrama la señalización en DL y por lo tanto reaccionar de forma instantánea ante mensajes de la red (retransmisiones en DL, asignaciones de transmisión en UL, etc.). Sin embargo, habitualmente el tráfico de paquetes es altamente variable, con periodos de transmisión dispersos y largos periodos de silencio. Mantener el receptor del terminal apagado durante esos periodos de silencio permite ahorrar energía de manera sensible en el terminal, lo que compensa ampliamente el coste en el incremento de la latencia.

LTE proporciona mecanismos para implementar DRX. De esta forma un terminal no está continuamente monitorizando el PDCCH en cada subtrama y se le permite pasar a un estado de ahorro de potencia (*power saving mode*) [10]. La DRX está controlada por la capa RRC. Básicamente, DRX establece un ciclo de operación básico por terminal (*DRX cycle*) cuya duración es configurable por la red. Dentro de cada *DRX cycle*, el terminal monitoriza el PDCCH sólo durante una subtrama para ver si tiene asignación de recursos, pasando al estado de inactividad el resto de subtramas. Evidentemente, cuanto más largo sea el *DRX cycle* más ahorro de energía se consigue, aunque a costa de una pérdida de rendimiento en términos de latencia⁹.

La Figura 6-16 ilustra el funcionamiento de DRX en LTE. Como se puede apreciar, existen diversos contadores (*timers*) que permiten controlar los periodos de actividad. El terminal se encuentra activo durante el *On Duration timer* que se inicia cada vez que llega la subtrama inicial de un *DRX cycle*. Si durante este tiempo el terminal recibe en una subtrama del PDCCH un mensaje nuevo de asignación de scheduling (DL o UL), se inicia un contador de tiempo de inactividad (*Inactivity Timer*), que evita que el terminal pase al estado de inactividad hasta que expire. De esta forma, si un terminal ha tenido actividad en una trama, se mantiene activo durante el *Inactivity Timer*, ya que es bastante probable que necesite transmitir o recibir más datos a corto plazo. Cada vez que se reciben mensajes de asignación, se reinicia el *Inactivity Timer*. Es importante destacar que la red sólo asignará nuevos recursos al terminal para cuando se encuentre activo, lo que supone una restricción adicional a los mecanismos de *scheduling* explicados anteriormente.

En cuanto a las retransmisiones híbridas (HARQ), éstas tienen lugar fuera de la operación DRX, de forma que si es necesaria alguna retransmisión no hay que esperar a los ciclos de DRX. Es decir, el terminal pasará al estado activo siempre que tenga que recibir/enviar un paquete de retransmisión, como se explica a continuación. Si el terminal no ha podido decodificar adecuadamente algún paquete durante el *On-duration timer*, entonces creará un nuevo proceso HARQ y pasará a la inactividad. Tal y como se vio en el Anexo 5.3 del Capítulo 5, en UL el mecanismo de HARQ es síncrono, de modo que tanto terminal como eNB conocen las subtramas en que el terminal reenviará los paquetes, pasando al estado activo de nuevo en esas subtramas. Por otra parte, en DL, el mecanismo HARQ es asíncrono, donde la subtrama donde se recibirá el paquete no está fijada. Por lo tanto, tras crear el proceso HARQ, el terminal iniciará un contador *HARQ Round Trip Time (RTT) Timer* asociado a dicho proceso antes de pasar a la inactividad. Este contador determina el tiempo mínimo que transcurre hasta que el terminal puede recibir una retransmisión por parte del eNB. Por lo tanto, cuando expire, el terminal se activará. Para evitar que la red esté forzada a enviar el paquete de retransmisión justo en la subtrama en que expira el *HARQ RTT Timer*, existe un segundo contador denominado *DRX Retransmission Timer* que se activa cuando expira el *HARQ RTT Timer* y que mantiene activo al terminal hasta la recepción del paquete retransmitido. Por lo tanto, el valor de

⁹ En este apartado se describe DRX cuando el terminal se encuentra en modo RRC-CONNECTED. En caso de que el terminal se encuentre en modo RRC-IDLE, es decir, sin radio bearers establecidos, el terminal monitoriza el canal de paging PCH cada *DRX cycle*.

este contador afecta al compromiso entre la flexibilidad a la hora de planificar retransmisiones y el consumo energético del terminal. Si cuando se recibe el paquete retransmitido sigue siendo imposible decodificarlo, se vuelve a activar el HARQ RTT Timer, volviéndose a repetir lo anterior. En caso contrario, se sigue el ciclo normal de DRX.

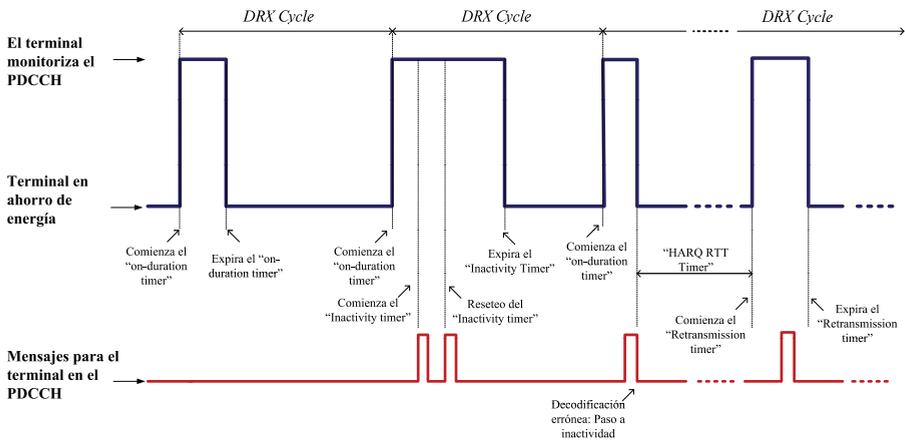


Figura 6-16 Mecanismo de recepción discontinua en LTE

El procedimiento descrito anteriormente sería suficiente para implementar satisfactoriamente DRX en LTE. No obstante y de forma opcional, se contempla la posibilidad de tener un ciclo de DRX corto (*Short DRX cycle*) adecuado para servicios como VoIP que se caracterizan por periodos cortos de actividad e inactividad más o menos en intervalos regulares. Si un terminal está configurado con *Short DRX cycle* sigue un ciclo de DRX corto (e.g., de 20 ms, que es el tiempo en el que el códec de VoIP entrega paquetes), donde al igual que con el ciclo largo también se implementa el mecanismo del contador de inactividad cada vez que existe actividad. Es importante destacar que un terminal puede pasar del ciclo corto al largo y viceversa. Para pasar del ciclo largo al corto, la red configura explícitamente al terminal. Para pasar del ciclo corto al largo existe un contador (*Short DRX Timer*) que se inicia cuando expira el *Inactivity Timer*, y que cuando expira hace que el terminal abandone la DRX con ciclo corto y pase a seguir el ciclo largo.

Conviene remarcar que una gestión inadecuada de la DRX podría afectar significativamente a la calidad de servicio del terminal, ya que afecta a la escucha del PDCCH donde se transmiten los mensajes de asignación de recursos tanto para DL como para UL. La gestión de recursos radio se debe encargar en este caso de ajustar adecuadamente la duración de los *timers* en función del tráfico del usuario y de su predictibilidad (por ejemplo, con el paso a *Short DRX cycle* si hay actividad vocal). Por otro lado, si la duración del *DRX cycle* es excesiva, puede que el terminal pierda el sincronismo, debiendo acceder al sistema a través del RACH. Por lo tanto, una gestión inadecuada del *DRX cycle* podría afectar a la capacidad de respuesta del sistema a través del RACH. Finalmente, la DRX puede afectar el desempeño de las estrategias de *scheduling*, ya que durante determinados periodos el terminal no se

encuentra disponible, pudiendo desperdiciarse oportunidades adecuadas para la transmisión y desaprovechando así la diversidad multiusuario.

6.10 Gestión flexible del espectro

La flexibilidad del espectro es una de las características clave de LTE. Ésta consiste en varios pasos que incluyen la flexibilidad en la banda de despliegue, la flexibilidad en el ancho de banda usado dentro de la banda de despliegue, y la flexibilidad propia de la interfaz radio basada en acceso OFDMA que permite una optimización de su uso por medio de técnicas de asignación dinámica de espectro DSA (*Dynamic Spectrum Assignment*). En este apartado se cubren cada uno de estos aspectos.

6.10.1 Flexibilidad en la banda de despliegue

La interfaz radio LTE es relativamente independiente de la banda de despliegue y en principio se puede desplegar en las bandas ya identificadas para IMT-2000 como en otras nuevas bandas que puedan aparecer en el futuro [34]. Tanto las bandas ya identificadas para LTE como futuras bandas pueden ser pareadas (para despliegues FDD: Duplexado por División en Frecuencia) o no pareadas (para despliegues TDD: Duplexado por División en Tiempo). La siguiente tabla muestra las bandas de operación que se han identificado hasta el momento para LTE.

Existen ciertas limitaciones en lo relativo a los requerimientos de radiofrecuencia asociados a una determinada banda. En concreto, estas restricciones se ven motivadas por:

- La coexistencia de los operadores en una misma zona geográfica. Estos operadores pueden desplegar distintas interfaces radio, y por lo tanto LTE debe especificar ciertos requerimientos que aseguren su coexistencia con otras tecnologías de acceso radio [35]. También afecta a despliegues donde distintos operadores comparten un mismo emplazamiento.
- La coexistencia de LTE con otros servicios o sistemas desplegados en bandas adyacentes o en la frontera del país, donde cambia el marco regulador del espectro radioeléctrico.
- El hecho de que LTE sea un sistema abierto desde el punto de vista de la banda de operación, de forma que futuras versiones del estándar agregan nuevas bandas de operación. Eso supone que los terminales se tengan que diseñar asumiendo que podrían funcionar en un amplio margen de frecuencias.

Estos requerimientos se suelen definir en términos de la máxima potencia permitida, las emisiones espúreas, emisiones fuera de banda (*Adjacent Channel Leakage Ratio*), etc.

Tabla 6.2 Bandas mejor identificadas para el despliegue de LTE en la Release [34]

Banda LTE	Banda para UL	Banda para DL	Tipo de Duplexado
1	1920 MHz – 1980 MHz	2110 MHz – 2170 MHz	FDD
2	1850 MHz – 1910 MHz	1930 MHz – 1990 MHz	FDD
3	1710 MHz – 1785 MHz	1805 MHz – 1880 MHz	FDD
4	1710 MHz – 1755 MHz	2110 MHz – 2155 MHz	FDD
5	824 MHz – 849 MHz	869 MHz – 894 MHz	FDD
6	830 MHz – 840 MHz	875 MHz – 885 MHz	FDD
7	2500 MHz – 2570 MHz	2620 MHz – 2690 MHz	FDD
8	880 MHz – 915 MHz	925 MHz – 960 MHz	FDD
9	1749.9 MHz – 1784.9 MHz	1844.9 MHz – 1879.9 MHz	FDD
10	1710 MHz – 1770 MHz	2110 MHz – 2170 MHz	FDD
11	1427.9 MHz – 1452.9 MHz	1475.9 MHz – 1500.9 MHz	FDD
12	698 MHz – 716 MHz	728 MHz – 746 MHz	FDD
13	777 MHz – 787 MHz	746 MHz – 756 MHz	FDD
14	788 MHz – 798 MHz	758 MHz – 768 MHz	FDD
17	704 MHz – 716 MHz	734 MHz – 746 MHz	FDD
33	1900 MHz – 1920 MHz	1900 MHz – 1920 MHz	TDD
34	2010 MHz – 2025 MHz	2010 MHz – 2025 MHz	TDD
35	1850 MHz – 1910 MHz	1850 MHz – 1910 MHz	TDD
36	1930 MHz – 1990 MHz	1930 MHz – 1990 MHz	TDD
37	1910 MHz – 1930 MHz	1910 MHz – 1930 MHz	TDD
38	2570 MHz – 2620 MHz	2570 MHz – 2620 MHz	TDD
39	1880 MHz – 1920 MHz	1880 MHz – 1920 MHz	TDD
40	2300 MHz – 2400 MHz	2300 MHz – 2400 MHz	TDD

6.10.2 Flexibilidad en el ancho de banda en uso

La interfaz radio de LTE basada en OFDMA ofrece una gran flexibilidad a la hora de desplegar el sistema dentro de una determinada banda, donde idealmente cualquier ancho de banda puede ser usado (en pasos de 180 kHz correspondientes a la anchura de banda de un PRB). No obstante, para facilitar la operación, LTE define un subconjunto de anchos de banda posibles. En concreto, se permiten bloques de espectro con 6, 15, 25, 50, 75 y 100 PRBs que se corresponden con anchos de banda nominales de 1.4 MHz, 3 MHz, 5 MHz, 10 MHz, 15 MHz y 20 MHz respectivamente [34]. Nótese que ancho de banda de estos bloques sería algo menor que el ancho de banda nominal debido a las bandas de guarda para reducir las emisiones fuera de banda. Conviene destacar que no todas las configuraciones de ancho de banda están disponibles para todas las bandas de operación definidas en el apartado anterior (véase la tabla 5.6.1.1 en [34]).

Esta flexibilidad en el ancho de banda en uso por parte de un operador permitirá, dentro del marco conocido como *spectrum refarming* (Capítulo 1), el despliegue gradual de LTE en bandas asignadas previamente a otros sistemas como GSM o UMTS. Por ejemplo, en un despliegue GSM de 15 MHz con radiocanales de 200 kHz sería posible introducir LTE en tres etapas, donde en cada etapa se despliega LTE en una porción de aproximadamente 5 MHz de radiocanales GSM, a medida que decrezca el tráfico cursado sobre GSM. Nótese que el *refarming* de LTE en bandas reservadas a sistemas previos se debe a que algunas de las bandas LTE presentadas en el apartado anterior coinciden dichas bandas.

6.10.3 Asignación dinámica de espectro

Las estrategias que se plantean para la planificación de frecuencia en LTE (FRF1, FRF3, etc.) que se vieron en el subapartado 6.5.1 están pensadas para distribuciones homogéneas del tráfico en el área de cobertura. Sin embargo, en el caso de distribuciones heterogéneas o variables, que en realidad son las que usualmente van a aparecer en escenarios de tipo práctico, estas estrategias se han mostrado claramente ineficientes [36], limitando la calidad de servicio en las celdas más cargadas o infrautilizando el espectro asignado en las menos cargadas. Por el contrario, la interfaz radio basada en OFDMA de LTE ofrece una excelente flexibilidad y abre un amplio abanico de posibilidades a la hora de gestionar el espectro en un determinado despliegue. Así, las estrategias de asignación dinámica de espectro plantean cómo conseguir una mejor eficiencia espectral en términos de número de usuarios por unidad de banda (o equivalentemente el *throughput* por célula y unidad de banda). Se introduce a este efecto una reutilización espacial variable de la banda asignada, a diferencia de lo que hasta ahora aparece en la literatura como esquemas fijos de asignación.

Como ejemplo, se citan dos metodologías para dichos algoritmos: algoritmos heurísticos [37] y algoritmos basados en *reinforcement learning* [38]. Los primeros son algoritmos que tratan de resolver el problema de forma intuitiva, esto es, se decide el número de frecuencias por cada celda en función de su carga para cumplir con unos objetivos de calidad de servicio y posteriormente se trata de repartir estos recursos de forma que se reduzca la interferencia intercelular. De esta manera, se mejora la eficiencia espectral cuando se compara con los esquemas clásicos (no dinámicos) de asignación de frecuencias. Por otro lado, los segundos algoritmos emplean la metodología de *reinforcement learning* para encontrar asignaciones de frecuencias por celda que tratan de optimizar un determinado indicador de recompensa. La metodología de *reinforcement learning* prueba distintas soluciones y aprende para converger a las mejores y descartar las peores en términos de recompensa. Por lo tanto, construyendo un indicador de recompensa que capture los objetivos de una asignación dinámica del espectro es posible aplicar esta metodología para encontrar asignaciones adecuadas.

La Figura 6-17 muestra una comparación en términos de incumplimiento de la calidad de servicio y eficiencia espectral entre los esquemas FRF1, FRF3 y las dos estrategias dinámicas descritas anteriormente. En estos resultados, se parte de una distribución homogénea del tráfico entre las celdas y a partir de los 30 minutos se pasa a una distribución heterogénea.

Observe como en este último caso, las estrategias dinámicas consiguen mantener la calidad de servicio en el sistema al mismo tiempo que se mantiene o mejora la eficiencia espectral.

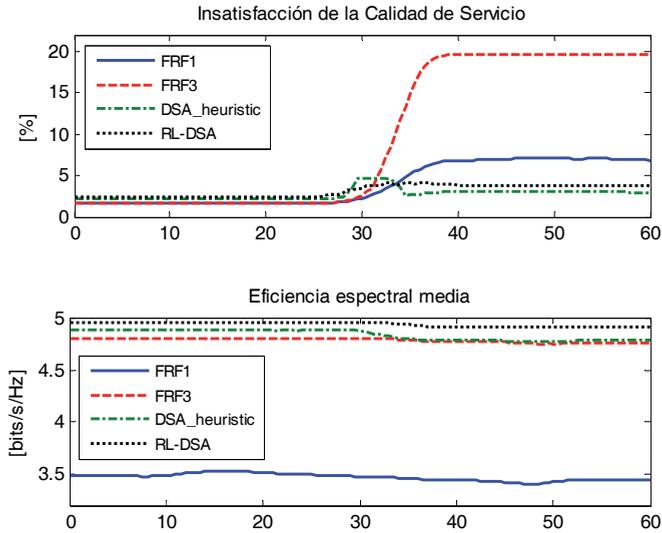


Figura 6-17 Comparación entre esquemas tradicionales de planificación y estrategias DSA

6.11 Referencias bibliográficas

- [1] J. Pérez-Romero, O. Sallent, R. Agustí, M.A. Díaz-Guerra, "Radio Resource Management Strategies in UMTS", John Wiley & Sons Ltd, England, 2005.
- [2] 3GPP TS 36.300 v8.11.0 "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2" enero 2010.
- [3] M. Anas, C. Rosa, F.D. Calabrese, P.H. Michaelsen, K.I. Pedersen, P.E. Mogensen, "QoS-Aware Single Cell Admission Control for UTRAN LTE Uplink," IEEE Vehicular Technology Conference, 2008. VTC Spring 2008. 11-14 mayo 2008 Page(s):2487 – 2491.
- [4] F.M. Sallabi, K. Shuaib, "Downlink call admission control algorithm with look-ahead calls for 3GPP LTE mobile networks" International Conference On Communications And Mobile Computing 2009, Leipzig, Alemania, pp: 712-715, 2009.
- [5] 3GPP TS 23.203 v.8.8.0, "Technical Specification Group Services and System Aspects, Policy and Charging control architecture", diciembre 2009.
- [6] 3GPP TS 23.401 v.8.8.0, "Technical Specification Group Services and System Aspects, General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access", diciembre 2009.

- [7] 3GPP TS 45.005 v8.5.0, "Technical Specification Group GSM/EDGE Radio Access Network; Radio transmission and reception" mayo 2009.
- [8] 3GPP TS 36.213 v.8.8.0, "Technical Specification Group Radio Access Network, Evolved Universal Terrestrial Radio Access (E-UTRA), Physical Layer Procedures", septiembre 2009.
- [9] R1-060048, "Channel Dependent Packet Scheduling for Single Carrier FDMA in E-UTRA Uplink", 3GPP TSG RAN WG1 LTE Ad Hoc Meeting, Helsinki, Finland, 23-25 enero 2006.
- [10] 3GPP TS 36.321 v.8.8.0, "Technical Specification Group Radio Access Network, Evolved Universal Terrestrial Radio Access (E-UTRA), Medium Access Control (MAC) protocol specifications", enero 2010.
- [11] C. Wengerter, J. Ohlhorst, A.G.E. von Elbwart, "Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA," IEEE 61st VTC 2005-Spring. Jun. 2005.
- [12] 3GPP TS 36.423 v8.8.0, "Technical Specification Group Radio Access Network, Evolved Universal Terrestrial Radio Access (E-UTRA), X2 application protocol (X2AP)", diciembre 2009.
- [13] A. Simonsson, "Frequency reuse and intercell interference coordination in E-UTRA". IEEE VTC Spring, 2007, 3091–3095.
- [14] S-E. Elayoubi, O. Ben Haddada, B. Fourestie, "Performance Evaluation of frequency Planning schemes in OFDMA-based Networks", IEEE Trans. Wireless Commun.; 7(5) pp. 1623-1633. 2008.
- [15] R1-080950, "Performance of proactive uplink ICIC under fractional load", Nokia Siemens Networks, Nokia. 3GPP TSG RAN WG1 #52, Sorrento (Italy), February, 2008.
- [16] R1-074444, "On Inter-Cell Interference Coordination Schemes without/with Traffic load Indication", Ericsson. 3GPP TSG RAN WG1 #50b, Shanghai (China), October, 2007.
- [17] G. Bourdreau, et.al. "Interference Coordination and Cancellation for 4G Networks", IEEE Communications Magazine, vol. 47, 4, abril 2009.
- [18] C. Castellanos et al., "Performance of Uplink fractional Power Control in UTRAN LTE", IEEE VTC Spring 2008, Singapore, pp. 2517-2521.
- [19] NGMN TE WP1, "Radio Performance Evaluation Phase 2 Report", Feb. 2008.
- [20] H. Hu et al. "Radio Resource Management for Cooperative Wireless Communication Systems with Organized Beam-Hopping Techniques", IEEE wireless commun. Vol. 15, 2, abril 2008, pp. 100-109.
- [21] Z. Bai, et al "On MIMO with successive interference cancellation applied to UTRA LTE" 3rd International Symposium on Communications, Control and Signal Processing,

2008. ISCCSP 2008. 12-14 pp. 1009-1013, marzo 2008.
- [22] L.G. Barbero, J.S. Thompson, "Performance of the Complex Sphere Decoder in Spatially Correlated MIMO channels", IET Commun., 2007, pp. 112 -130.
 - [23] W. Choi, J. Andrews, "The Capacity Gain from bas Station Cooperative Scheduling in a MIMO DPC Cellular System," ISIT 2006, Seattle, pp. 1224-1228.
 - [24] 3GPP TS 36.214 v8.7.0, "Technical Specification Group Radio Access Network, Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer - Measurements", septiembre 2009.
 - [25] 3GPP TS 36.304 v8.8.0, "Technical Specification Group Radio Access Network, Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) procedures in idle mode", enero 2010.
 - [26] 3GPP TS 36.133 v8.8.0, "Technical Specification Group Radio Access Network, Evolved Universal Terrestrial Radio Access (E-UTRA); Requirements for support of radio resource management", diciembre 2009.
 - [27] 3GPP TS 36.331 v8.8.0, "Technical Specification Group Radio Access Network, Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification", enero 2010.
 - [28] T.-H. Kim, Q. Yang, J.-H. Lee, S.-G. Park, Y.-S. Shin "A Mobility Management Technique with Simple Handover Prediction for 3G LTE Systems", IEEE 66th Vehicular Technology Conference Fall, 2007; Baltimore, MD 2007.
 - [29] M. Anas, F.D. Calabrese, P.E. Mogensen, C.Rosa, K.I. Pedersen, "Performance Evaluation of Received Signal Strength Based Hard Handover for UTRAN LTE", IEEE 65th Vehicular Technology Conference, abril 2007.
 - [30] M. Anas, F.D. Calabrese, P.-E.Ostling, K.I. Pedersen, .E. Mogensen, "Performance Anayl-sis of Handover Measurements and Layer-3 Filtering for UTRAN LTE",. IEEE 18th International Symposium on In Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007.
 - [31] D. Aziz, R. Sigle, "Improvement of LTE Handover Performance through Interference Co-ordination" IEEE 69th Vehicular Technology Conference, 2009. VTC Spring 2009. 26-29 abril 2009.
 - [32] S. Jeux, G. Mange, P. Arnold, F. Bernardo (editores), "Simulation based recommendations for DSA and self-management" E3 Deliverable D3.3. <https://ict-e3.eu/project/deliverables/deliverables.html>.
 - [33] W. Ju, W.M. Cioffi, "On constant power water filling" IEEE international conference on communications (ICC), 2001.
 - [34] 3GPP TS36.101 v8.8.0 "Technical Specification Group Radio Access Network; Evolved

Universal Terrestrial Radio Access (E-UTRA) User Equipment (UE) radio transmission and reception”, diciembre 2009.

- [35] M.H. Ng, S.D. Lin, J. Li, S. Tatesh, “Coexistence Studies for 3GPP LTE with Other Mobile Systems” IEEE Comm. Magazine, vol 47, issue 4, pp. 60-65, abril 2009.
- [36] D. López-Pérez, A. Jüttner, and J. Zhang, “Dynamic Frequency Planning Versus Frequency Re-use Schemes in OFDMA Networks” IEEE VTC (Vehicular Technology Conference), Barcelona, Spain, 26-29 abril 2009.
- [37] F. Bernardo , R. Agustí, J. Pérez-Romero, O. Sallent “Dynamic Spectrum Assignment in Multicell OFDMA Networks enabling a Secondary Spectrum Usage”, Wireless Communications and Mobile Computing , abril 2009.
- [38] F. Bernardo, R. Agustí, J. Pérez-Romero, and O. Sallent “A Novel framework for Dynamic Spectrum Assignment in MultiCell OFDMA Networks based on Reinforcement Learning”, Wireless Communications and Networking Conference (WCNC2009), Budapest, Hungary, 5-8 abril. 2009.



VII

Redes Heterogéneas

7 Redes Heterogéneas

7.1 Generalidades

La convivencia de redes de acceso radio de diferentes tecnologías constituye una característica inherente a la propia evolución de los sistemas de comunicaciones móviles. La adopción de nuevas tecnologías radio pocas veces representa un proceso disruptivo sino que resulta habitual que diferentes soluciones formen parte de una misma infraestructura de red. Los motivos son claros: la introducción de nuevas tecnologías, o mejoras sustanciales en las tecnologías existentes, requiere un importante esfuerzo económico para su despliegue a la vez que resulta necesario rentabilizar al máximo las inversiones realizadas en la infraestructura existente. Por ello, es común encontrar operadores móviles que ofrecen sus servicios a través de, por ejemplo, GSM, UMTS R99 y UMTS HSPA, y en un futuro, LTE. Por otro lado, existen soluciones tecnológicas que son en cierta medida complementarias en cuanto a su idoneidad para ser utilizadas en determinados entornos de operación (e.g., tecnologías optimizadas para proporcionar acceso en movilidad en exteriores o bien optimizadas para operar en interiores de edificios). Por este motivo, la convivencia de múltiples y diversas tecnologías también viene propiciada por su adecuación a diferentes aplicaciones y escenarios de operación. A modo de ejemplo, las soluciones Wi-Fi ampliamente utilizadas para proporcionar acceso radio a redes fijas de comunicaciones de banda ancha, representan una alternativa interesante para la provisión de los servicios propios de las redes celulares en entornos domésticos, oficinas, hoteles, aeropuertos, etc. Por tanto, la propia evolución y mejora de los sistemas de comunicaciones, junto con la utilización de las tecnologías más adecuadas en cada entorno de operación, conduce a la coexistencia de diferentes tecnologías radio en los sistemas de comunicaciones móviles. Las redes de comunicaciones resultantes donde se utilizan diferentes tecnologías radio para la provisión de los servicios se conocen como redes heterogéneas.

Asociada a la existencia de redes heterogéneas aparece la necesidad de disponer de terminales multimodo, es decir, terminales que sean capaces de conectarse a la red mediante la utilización de varias tecnologías de acceso radio. En este contexto, es importante destacar que la mayoría de terminales 3G/UMTS que hay actualmente en el mercado son a la vez terminales 2G/GSM, y, sin lugar a dudas, muchos de los futuros terminales LTE van a ser terminales multimodo que soporten algunas de las tecnologías actuales (e.g., *dongles* para acceso de banda ancha móvil mediante HSPA y LTE).

El sistema LTE ha sido específicamente diseñado para soportar despliegues de red donde la red de acceso E-UTRAN co-exista con otras redes de acceso radio. En particular, el servicio de conectividad IP de la red troncal EPC puede ofrecerse también a través de las redes de acceso 3GPP anteriores a LTE, es decir, UTRAN y GERAN. La interconexión de estas redes de acceso a una red LTE se realiza a través de interfaces específicas con la nueva red troncal EPC de forma que se garantiza la continuidad de los servicios cuando los terminales cambian de red de acceso (i.e., se soportan mecanismos de *handover* entre tecnologías 3GPP). Además, el sistema LTE contempla también la posibilidad de utilizar redes de acceso no estandariza-

das por 3GPP, redes que denominaremos en adelante simplemente como redes no 3GPP, para la provisión de los servicios de la red troncal EPC. Dentro de las redes no 3GPP, cabe mencionar especialmente las redes CDMA2000 desarrolladas por 3GPP2 (e.g., 1xRTT y EV-DO), para las cuales el sistema LTE se ha incorporado de forma clara en la hoja de ruta de su proceso evolutivo [1].

En el apartado 7.2 de este capítulo se describen los fundamentos de las redes heterogéneas y se analizan de forma genérica los principales mecanismos de *interworking* que permiten la co-existencia y explotación conjunta de diferentes tecnologías de acceso radio. Los conceptos que se desarrollan en el apartado 7.2 constituyen un marco de referencia útil para la comprensión y el análisis de las diferentes soluciones de *interworking* que han sido especificadas en el sistema LTE y que son objeto de estudio en los apartados posteriores. Así, en el apartado 7.3 se abordan las soluciones de *interworking* entre LTE y las redes de acceso 3GPP anteriores a E-UTRAN y, finalmente, en el apartado 7.4 se describe el soporte que ofrece el sistema LTE en cuanto a la utilización de redes de acceso alternativas no 3GPP.

7.2 Fundamentos de las soluciones de interworking

El término *interworking* se utiliza en este capítulo para hacer referencia al conjunto de funciones y mecanismos mediante los cuales se consigue un determinado nivel de interacción (i.e., nivel de *interworking*) en la operación de dos o más redes de comunicaciones. El nivel de *interworking* requerido entre redes admite una amplia gradación atendiendo a los objetivos perseguidos a través de la operación conjunta de las redes. Así, pueden plantearse soluciones de *interworking* básicamente concebidas para la provisión de servicios en itinerancia, hasta soluciones que persigan la continuidad de servicios entre redes de forma transparente al usuario (*seamless mobility*) y que permitan una gestión común o coordinada de los recursos de comunicaciones de ambas redes.

A continuación se describe un escenario de referencia a partir del cual puede analizarse el alcance y necesidad de las soluciones de *interworking* que puedan plantearse en el contexto de redes heterogéneas. Posteriormente, en los apartados 7.2.2 y 7.2.3, se realizan un conjunto de consideraciones básicas sobre la arquitectura de las redes de acceso y terminales en aras a identificar aquellos aspectos clave que inciden en el diseño de mecanismos de *interworking*. En base a tales consideraciones, en los siguientes apartados se realiza un análisis amplio y estructurado de los diferentes niveles de *interworking* (apartado 7.2.4) y de los mecanismos de *interworking* pertinentes en cada nivel (apartado 7.2.5).

7.2.1 Escenario de referencia

En la Figura 7.1 se ilustra el escenario de referencia considerado en el análisis de los niveles de *interworking* entre redes. El escenario de referencia considera dos redes A y B, cada una basada en una tecnología radio diferente, RAT A y RAT B, que pueden tener coberturas solapadas en determinadas ubicaciones de la zona de servicio. El grado de solapamiento de la

cobertura puede ser muy diverso. Así, cabría plantearse tanto escenarios donde ambas redes están disponibles en toda la zona de servicio (e.g., en una ciudad un operador ofrece tanto cobertura UMTS como GSM) hasta escenarios donde el solape únicamente fuera puntual en determinadas zonas (e.g., en un futuro despliegue de LTE, la cobertura de E-UTRAN se limita a un número reducido de emplazamientos a la vez que se garantiza la cobertura completa del área de servicio mediante UMTS).

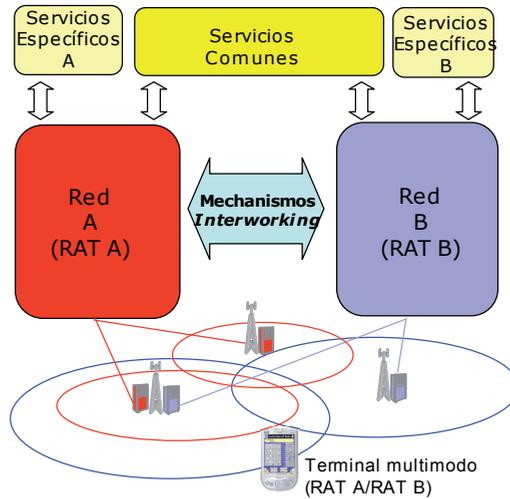


Figura 7.1 Escenario de referencia para el planteamiento de soluciones de interworking en redes heterogéneas

En cuanto a la provisión de servicios, tal como se muestra en la Figura 7.1, se considera que ambas redes pueden proporcionar acceso a un conjunto común de servicios (e.g., el servicio de acceso a Internet podría ofrecerse tanto a través de la red A como de la red B). Asimismo, podría suceder que cada red proporcionara, de forma exclusiva, acceso a una serie de servicios específicos. Por ejemplo, en el caso de LTE y UMTS, la red LTE podría utilizarse para proporcionar un servicio de video bajo demanda de alta calidad que no sería accesible mediante UMTS, y en cambio, el servicio telefónico podría ofrecerse únicamente a través de UMTS.

Finalmente, en cuanto a los derechos de acceso a cada una de las redes, un caso general sería el que las redes A y B pertenecieran a operadores diferentes y el acceso de los usuarios de un operador a través de la red del otro estuviera sujeto a los acuerdos de itinerancia pertinentes entre operadores. Por ello, en aras a cubrir este caso general, en la descripción de los niveles de *interworking* se utiliza la notación red matriz y red visitada para indicar, respectivamente, la red a la que se encuentra suscrito un usuario y la red a la que puede conectarse un usuario sin tener una suscripción en ella. Nótese que un caso más particular que también puede derivarse del escenario genérico mostrado en la Figura 7.1 sería el de un único ope-

rador de red que utiliza dos redes de acceso diferentes (e.g., GSM y UMTS) de forma que la suscripción de un usuario es válida en ambas.

Atendiendo al escenario de referencia planteado, la provisión de acceso a los diferentes servicios disponibles (comunes y/o específicos) a través de cualquiera de las redes mediante el uso de terminales multimodo requiere de la existencia de los mecanismos de *interworking* apropiados. Los mecanismos de *interworking* tendrán un papel clave en la gestión de la movilidad entre redes y en la explotación de mecanismos de gestión de recursos radio que contemplen de forma conjunta ambos accesos.

7.2.2 Arquitectura genérica de las redes de acceso

Respecto a la arquitectura de red, tal como se ha analizado en el Capítulo 2, las futuras redes de acceso radio están principalmente diseñadas para proporcionar servicios de conectividad de paquetes con un determinado perfil de calidad de servicio (QoS). Sobre esta base, la provisión de servicios finales se realiza a través de plataformas especializadas externas a la red de acceso (e.g., *IP Multimedia Subsystem*, IMS) que están accesibles a través del servicio de conectividad proporcionado.

En este contexto, en la Figura 7.2 se muestra una representación simplificada de los principales nodos que componen una arquitectura de red genérica representativa de la mayoría de los sistemas actuales de comunicaciones móviles basados en conmutación de paquetes (e.g., GPRS/UMTS, LTE, CDMA2000 1xEV-DO, Mobile WiMAX y redes de acceso WLAN). Junto con los nodos de la red de acceso, la figura muestra la ubicación en la red de las diferentes torres de protocolos con las que interactúa el terminal para la obtención del servicio de conectividad.

Tal como se observa en la Figura 7.2, la provisión del servicio de conectividad con redes externas y plataformas de servicios requiere de la utilización de pasarelas de red en la red de acceso. La pasarela de red es el nodo que alberga las principales funciones asociadas a la provisión del servicio de conectividad (i.e., terminación de los protocolos de la capa de red utilizada y funciones asociadas tales como, e.g., asignación de direcciones IP). Todo el tráfico cursado por la red de acceso fluye a través de una o múltiples pasarelas de red. Por ello, la pasarela de red constituye un elemento clave a través del cual pueden articularse algunas políticas de calidad de servicio (e.g., utilización de limitadores de tasa de los flujos de usuario que atraviesan la pasarela) y mecanismos de tarificación (e.g., recogida del volumen de bytes que cursa un usuario para su posterior envío a los sistemas de facturación). Todas estas funcionalidades de la pasarela de red son prácticamente independientes del tipo de tecnología radio utilizada. Ejemplos de pasarelas de red en las redes actuales serían el nodo P-GW en redes LTE y el nodo GGSN en redes GPRS/UMTS.

Por otro lado, la transmisión en la interfaz radio se sustenta en torres de protocolos que dependen de la tecnología radio utilizada. La terminación de los protocolos radio en la red puede ubicarse por completo en las estaciones base o puntos de acceso (como es el caso de

E-UTRAN y tecnologías IEEE 802) o bien implementarse de forma distribuida según un diseño jerárquico entre estaciones base y controladores radio (como es el caso de UTRAN y GERAN). La torre de protocolos radio comprende básicamente una capa física y una capa de enlace, y proporciona un servicio de transferencia de paquetes a través de la interfaz radio (e.g., paquetes IP en LTE, paquetes IP/PPP/X.25 en GPRS, paquetes IP y tramas Ethernet en IEEE 802.16, etc.). El servicio de transferencia de paquetes de la interfaz radio incluye funciones que permiten la aplicación de un trato específico de QoS (e.g., servicios de tasa garantizada, diferenciación de tráfico en base a prioridades en momentos de congestión, etc.). Las funcionalidades soportadas en la capa radio son claves para sustentar una gestión eficiente de los recursos radio (e.g., asignación dinámica de los recursos radio, control de la movilidad entre estaciones base, etc.).

Así pues, estaciones base y pasarelas de red, en sus diferentes variantes, constituyen los dos elementos fundamentales que forman parte de cualquier red de acceso en cuanto a la realización de las funciones necesarias para la transferencia del tráfico de los usuarios. Dichas funciones forman el denominado plano de usuario. Adicionalmente, junto con las estaciones base y pasarelas de red, el plano de usuario puede también incluir otros elementos de red (no mostrados en la figura) encargados de albergar, por ejemplo, funciones de anclaje necesarias en determinados esquemas de movilidad (e.g., pasarela S-GW en LTE, SGSN en GPRS/UMTS, nodo ASN, Access Service Network, en redes Mobile WiMAX, etc.). Dichas funciones de anclaje canalizarían la transferencia de paquetes de un determinado usuario entre las estaciones base y la pasarela de red permitiendo, por ejemplo, que un cambio de estación base, fuera transparente para la pasarela de red.

La gestión de las funciones asociadas al plano de usuario de la red de acceso se consigue a través del denominado plano de control. A diferencia del plano de usuario, las funciones del plano de control serían todas aquellas que no se encargan del procesado de los flujos de datos de los usuarios. En otras palabras, el plano de control de la red de acceso alberga las funciones necesarias para realizar el control de acceso a la red (e.g., autenticación y autorización), la gestión de movilidad (e.g., registro, actualización de la localización y aviso), la gestión de seguridad (e.g., gestión de claves y control de los servicios de confidencialidad e integridad) y la gestión de las sesiones (e.g., activación de los servicios de conectividad de la red). El soporte de todas estas funciones asociadas al plano de control de la red se esquematiza en la Figura 7.2 mediante la representación de un nodo de control y la torre de protocolos asociada a través de la cual la red interactúa con el terminal (e.g., el nodo de control en LTE sería la entidad MME, en redes GPRS/UMTS dichas funciones se integran en el nodo SGSN, etc.). Junto con el nodo de control, la red de acceso dispone también de bases de datos necesarias para sustentar la operativa de las funciones de control (e.g., almacenamiento de datos de localización, parámetros de seguridad, perfiles de QoS, etc.). Las funciones y torres de protocolos utilizadas en el plano de control dependen en gran medida de la tecnología de acceso radio (e.g., en LTE, la señalización de control entre la red y el terminal se lleva a cabo mediante los protocolos NAS diseñados por 3GPP a tal efecto).

Finalmente, la interconexión entre elementos de la red (nodos de control, estaciones base y pasarelas de red) se lleva a cabo básicamente mediante el uso de redes de transporte IP/ MPLS o Ethernet, tecnologías completamente independientes de la tecnología de acceso radio. Un ejemplo claro de esta tendencia es el propio sistema LTE donde, tal como se ha visto detalladamente en el Capítulo 2, todas las interfaces de la red LTE se soportan sobre torres de protocolos IP.

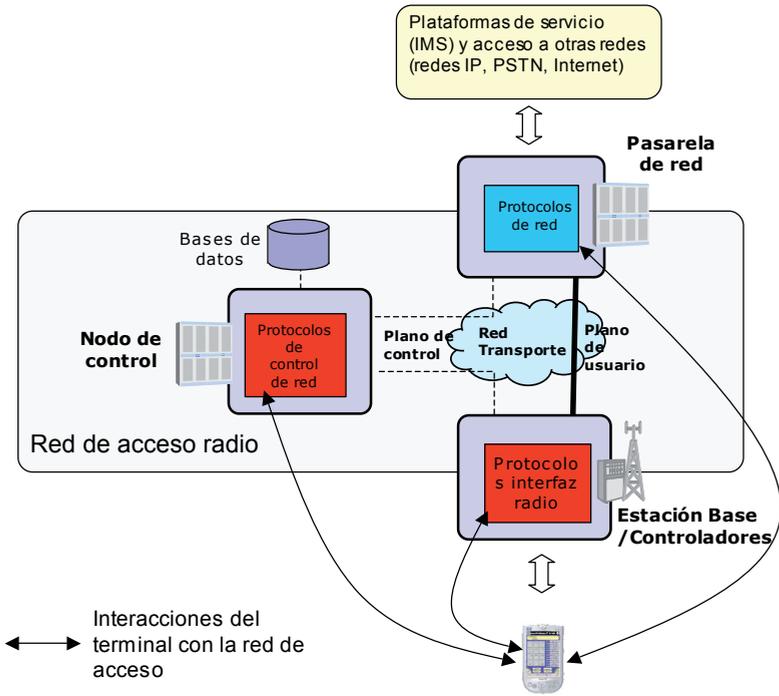


Figura 7.2 Arquitectura genérica de una red de acceso radio

7.2.3 Terminales multimodo

En la Figura 7.3 se muestra la torre de protocolos simplificada de lo que podría ser un terminal multimodo para redes de acceso como las descritas anteriormente. La torre de protocolos consiste en un conjunto de torres de protocolos propias de cada una de las tecnologías radio soportadas y un conjunto de protocolos comunes para las capas de red y superiores (i.e., transporte y aplicación).

En relación a los protocolos específicos de cada tecnología radio, el terminal incluye el soporte tanto de los protocolos propios de la interfaz radio (e.g., capas física y de enlace) como de los protocolos de control de red (e.g., protocolos NAS en redes 3GPP, protocolos EAP para autenticación en redes IEEE 802, etc.).

En cuanto a los protocolos comunes, la capa de red (e.g., IP) proporciona un sustrato uniforme sobre el cual pueden funcionar diferentes protocolos de la capa de transporte (e.g., TCP y UDP) y de aplicación (e.g., protocolo SIP) de forma independiente a la tecnología de acceso radio. Por ello, es importante destacar el papel fundamental que juega la capa de red en el modelo de *interworking* debido al hecho de que esta capa constituye el nexo natural entre las múltiples interfaces de comunicaciones que pueda tener el dispositivo multimodo y las capas de protocolos superiores.

Además de las consideraciones relativas a la estructura de la torre de protocolos, un aspecto clave de un terminal multimodo es el soporte de funciones asociadas a la gestión de movilidad entre redes. En particular, un terminal multimodo podría albergar una función encargada de seleccionar la mejor red de acceso en cada momento, atendiendo a la disponibilidad de la misma y, por ejemplo, al tipo de servicios cursados y preferencias del usuario. Dicha función de selección de red podría sustentarse en la información proporcionada por las diferentes capas radio (e.g., niveles de señal en cada interfaz) e iniciar los procedimientos de cambio de red cuando fuera necesario (e.g., mecanismos de *handover* iniciados por el terminal). Una alternativa o complemento a la gestión de la selección de red en el terminal multimodo es que las redes de acceso dispongan de mecanismos de control de la movilidad desde la red (e.g., iniciación de mecanismos de *handover* desde la red de acceso). En este caso, el terminal podría albergar funciones que asistan a las redes de acceso en la toma de decisiones de movilidad entre redes (e.g., envío de medidas de las redes disponibles desde el terminal).

Por último, una característica de los terminales multimodo que condiciona en gran medida la necesidad de incorporar determinados mecanismos de *interworking* es la capacidad del terminal para operar simultáneamente las dos interfaces radio. Permitir la operación simultánea puede incrementar enormemente la complejidad del diseño de los terminales para tener en cuenta los posibles acoplamientos en RF cuando las frecuencias utilizadas en las dos tecnologías radio estuvieran relativamente cercanas. Así, de forma general, los terminales multimodo pueden clasificarse en dos categorías: terminales multimodo dual-radio, que permiten mantener operativas las dos interfaces radio a la vez, y terminales multimodo single-radio, que solamente soportan la transmisión/recepción por una interfaz de forma simultánea. Tal como se verá más adelante, el soporte de mecanismos de *handover* para el caso de terminales single-radio es mucho más complejo si se quieren conseguir tiempos de interrupción del servicio reducidos ya que, inevitablemente, el terminal debe desactivar una interfaz y activar la otra en un cambio de red. Por el contrario, un terminal dual-radio podría estar conectado de forma simultánea a las dos redes de acceso de forma que el tiempo de interrupción del servicio podría ser mínimo (debido únicamente a procesos de cambios de encaminamientos dentro de las redes).

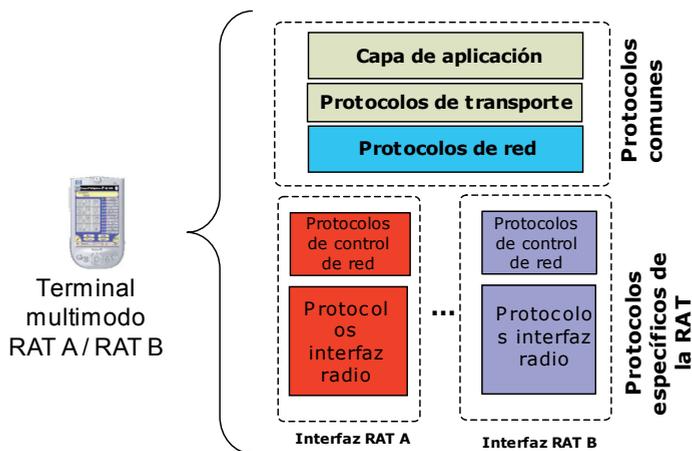


Figura 7.3 Arquitectura de protocolos de un terminal multimodo

7.2.4 Niveles de interworking

En este apartado se definen diferentes niveles de *interworking* que permiten establecer los requerimientos básicos que condicionan la elección y/o el diseño del tipo de mecanismos de *interworking* necesarios.

La definición de niveles *de interworking* puede realizarse en base a un planteamiento centrado en caracterizar la disponibilidad de acceso a los servicios por parte de los usuarios de la red. Este planteamiento permite desacoplar la definición de los niveles de *interworking* de aspectos particulares de arquitectura, tecnología y funciones de la red. Bajo este enfoque, se definen los siguientes cuatro niveles de *interworking* entre redes:

- **Nivel A: Acceso a servicios a través de una red visitada.** Este nivel de *interworking* permite a un usuario acceder a los servicios disponibles en una red visitada mediante el uso de la identidad y credenciales otorgadas por otra red (e.g, la red matriz donde el usuario tiene la suscripción). Un ejemplo sería el caso de un usuario de una red UMTS/LTE, equipado con un terminal portátil con interfaces de red celular UMTS/LTE y Wi-Fi, que pudiera obtener conectividad a Internet a través de una red Wi-Fi de acceso público (e.g., en un centro de conferencias, estación de tren, aeropuerto) utilizando como credenciales las disponibles en la tarjeta de identificación del usuario UMTS/LTE (USIM).
- **Nivel B: Acceso a servicios de la red matriz desde una red visitada.** Bajo este nivel, un usuario que se conecta a través de una red visitada podría acceder a servicios disponibles en la red de su operador matriz. Como ejemplo, retomando el caso del usuario con un terminal portátil con interfaces UMTS/LTE y Wi-Fi, el usuario podría acceder a los servicios UMTS/LTE (e.g., servicios de presencia o mensajería instantánea entre un

grupo cerrado de usuarios) mientras se encuentra conectado a través de la red Wi-Fi. Es importante destacar que ni el nivel A ni el B consideran aspectos de continuidad de servicio si el usuario cambiara de red durante el transcurso de sus conexiones.

- **Nivel C: Continuidad de servicio entre redes.** Este nivel extiende los dos niveles anteriores de forma que el usuario no tendría que re-establecer sus servicios al producirse un cambio en la red de acceso. Sin embargo, en este nivel se admite la posibilidad de que la calidad de servicio pueda sufrir un cierto deterioro durante la transición entre redes debido al propio mecanismo de cambio de red. Así pues, siguiendo con el ejemplo del usuario con el terminal portátil UMTS/LTE y Wi-Fi, el usuario podría encontrarse descargando un fichero de datos a través del acceso Wi-Fi mientras se está desplazando y pierde la cobertura de esa red. En ese momento, automáticamente el terminal conmuta a la red UMTS/LTE que proporciona cobertura en exteriores y la descarga del fichero continua a través de la nueva red, a pesar de que en la transición hubiera podido producirse una breve parada de la descarga.
- **Nivel D: Continuidad transparente de servicio entre redes.** Este nivel de *interworking* entre redes permitiría la continuidad de los servicios sin que se produjera una degradación perceptible al usuario durante la transición entre redes. Una forma de referirse a la continuidad transparente de servicio es también mediante el concepto de *seamless mobility*. La consecución de este nivel de *interworking* se fundamenta principalmente en el soporte de mecanismos de *handover* entre redes heterogéneas que permitan realizar el cambio de las conexiones entre redes con tiempos de interrupción y pérdidas de información muy reducidos (e.g., el tiempo de interrupción requerido en un proceso de *handover* entre LTE y UMTS debe ser inferior a 300 ms para servicios en tiempo real [1]). Según el ejemplo anterior, si el usuario con un terminal UMTS/LTE y Wi-Fi iniciará una conexión de voz sobre IP (VoIP) estando conectado a través de Wi-Fi, dicha conexión debería transferirse de forma transparente a la red LTE/UMTS si el usuario sale de la cobertura de la red de área local. Finalmente, es importante destacar en este punto que la capacidad de proporcionar una continuidad transparente de servicio depende también de la capacidad que tienen las redes involucradas de poder proporcionar el mismo servicio con igual nivel de QoS.

7.2.5 Mecanismos de interworking

En este apartado se describen los mecanismos de *interworking* más relevantes sobre los que se sustentan los niveles de *interworking* descritos en el apartado 7.2.4. El alcance de dichos mecanismos de *interworking* y su relación con las principales capas de protocolos de una red radio (descritas en el apartado 7.2.2) se ilustran en la Figura 7.4.

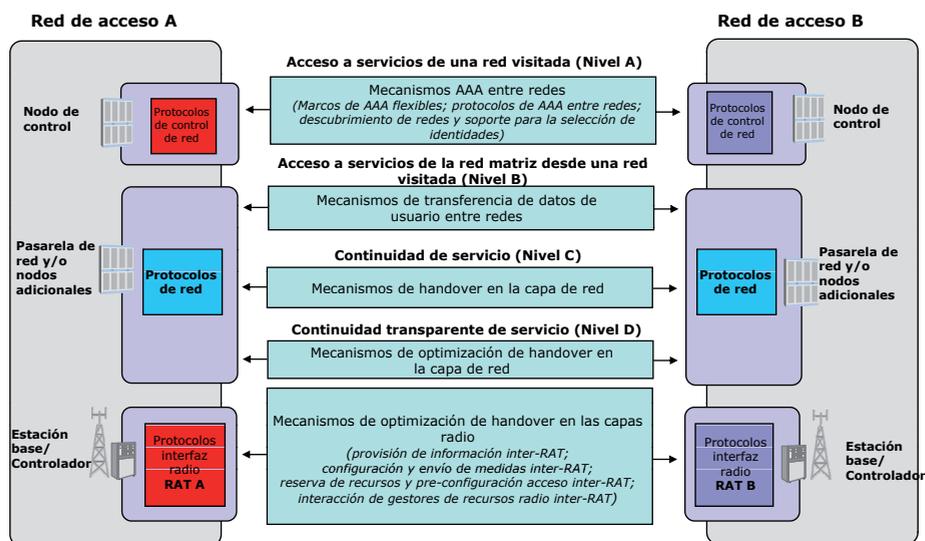


Figura 7.4 Niveles de interworking y mecanismos de interworking asociados

A continuación se describen los diferentes mecanismos de *interworking* mencionados en la Figura 7.4.

7.2.5.1 Mecanismos de AAA entre redes

La finalidad básica de los mecanismos que permiten la realización de un Nivel A es que un usuario pueda hacer uso de las credenciales otorgadas por su red matriz (con la que mantiene una suscripción) para obtener acceso a los servicios ofrecidos en una red visitada, no siendo necesario que la red visitada tenga ningún conocimiento de dicho usuario previamente a la solicitud del servicio. Asimismo, estos mecanismos deben permitir transferir información necesaria para la tarificación del acceso a los servicios a través de redes visitadas (e.g., la utilización de los recursos que el usuario hace en la red visitada se envía a la red matriz que se encarga de la facturación de sus suscriptores). Estos mecanismos se denominan de forma genérica como mecanismos AAA (*Authentication Authorization, and Accounting*). Entre los mecanismos de AAA principales para la realización del Nivel A se encuentran:

- Mecanismos de autenticación y autorización flexibles que permitan soportar múltiples métodos de autenticación en una misma red. Un ejemplo claro lo constituye el protocolo *Extensible Authentication Protocol* (EAP) definido en IETF RFC 3748 [3] que proporciona el soporte básico para el transporte fiable de diferentes protocolos de autenticación. A través del protocolo EAP se permite la autenticación del usuario mediante credenciales simples del tipo usuario/password así como mediante la utilización de certificados digitales. Marcos de autenticación y autorización basados en EAP han sido adoptados en el sistema LTE para el acceso genérico a través de redes no 3GPP (tal como se verá en

el apartado 7.4). EAP también forma parte del sistema de control de acceso soportado en equipos Wi-Fi que soporten *Wireless Protected Access (WPA)*, y en redes Mobile WiMAX. Ejemplos de protocolos de autenticación soportados sobre EAP son EAP-SIM, EAP-AKA y EAP-AKA¹, desarrollados para el soporte de los mecanismos de autenticación 3GPP en redes no 3GPP, y EAP-TTLS (*EAP Tunneled Transport Layer Security*) y PEAP (*Protected EAP*) ampliamente utilizados en instalaciones Wi-Fi.

- Funciones de *Proxy/Relay* para la transferencia de información AAA entre redes, junto con los correspondientes protocolos de señalización. En este contexto, el protocolo *Diameter* definido en la RFC 3588 [4] está concebido para poder cumplir los requisitos mínimos de un protocolo de AAA entre redes. Tal como se verá más adelante, el sistema LTE incorpora el uso de este protocolo en las soluciones de *interworking* con redes no 3GPP.
- Mecanismos mejorados para el descubrimiento de redes. Mediante estos mecanismos se pretende que las redes difundan, o den a conocer bajo demanda, las posibilidades que la red ofrece en términos de aceptar conexiones de usuarios de otras redes. En este sentido, las soluciones existentes en la actualidad para que un usuario conozca si una determinada red puede darle servicio, pasa por la disponibilidad de listas de redes pre-configuradas en los terminales o en tarjetas inteligentes de los usuarios donde se enumeran las redes potencialmente accesibles. Sin embargo, la proliferación de redes y operadores hace que los mecanismos basados en listas pre-configuradas no sean escalables de forma que es necesario disponer de soluciones más flexibles que de forma dinámica puedan dar a conocer dicha información de forma eficiente. A modo de ejemplo, la extensión 802.11u [5] pretende posibilitar que un terminal Wi-Fi pueda interactuar con la red sin necesidad de estar asociado a ella en aras a descubrir si sus credenciales pueden ser utilizadas en dicha red.

7.2.5.2 Mecanismos para la transferencia de datos de usuario entre redes

Estos mecanismos permiten la transferencia entre redes de la información generada o destinada a un usuario de forma que el usuario puede tener acceso a un determinado servicio ofrecido en una red diferente de la que se encuentre conectado. Estos mecanismos constituyen la base del Nivel B.

En redes de paquetes, los mecanismos habituales para conseguir dicho objetivo se basan en la utilización de protocolos como *Layer 2 Tunneling Protocol (L2TP)*, definido en RFC 2661 [6], y el modo túnel del protocolo *IPsec* definido en RFC 4301 [7]. Los túneles pueden establecerse de forma directa entre los terminales y elementos de red ubicados en una red remota desde donde se accede a los servicios o bien entre elementos de ambas redes de forma que el proceso se mantiene transparente al usuario. A modo de ejemplo, las especificaciones 3GPP para acceso mediante WLAN a servicios UMTS, establece la necesidad de que

¹ EAP-AKA¹ definido en la RFC 5448 es una versión mejorada del protocolo EAP-AKA especificado en RFC 4187.

los terminales utilicen el protocolo *IPsec Encapsulating Security Payload* (ESP) descrito en RFC 4303 [8] para establecer un túnel entre el propio terminal y un determinado nodo de la red UMTS. Mediante este túnel, un terminal conectado en una red Wi-Fi podría acceder a los servicios ofrecidos en la red UMTS de su operador celular (e.g., servicios en plataformas IMS).

7.2.5.3 Mecanismos de handover en la capa de red

El soporte de mecanismos de *handover* en la capa de red es necesario cuando la continuidad del servicio entre redes requiere que la dirección de red utilizada por el usuario (e.g., dirección IP) sobreviva al cambio de red de acceso. El requerimiento de mantener la dirección IP aplica, por ejemplo, a las conexiones realizadas con el protocolo TCP donde no se soporta que los nodos participantes cambien su dirección IP durante el transcurso de una conexión (i.e., la conexión TCP se corta si uno de los terminales deja de estar accesible a través de la dirección IP que utilizó durante su establecimiento). Los mecanismos de *handover* en la capa de red constituyen la base para la consecución de un Nivel C de *interworking*.

En este contexto, el estándar Mobile IPv4 (MIPv4) recogido en RFC 3344 [9] fue una de las primeras soluciones destinadas a cubrir las necesidades de itinerancia de usuarios que querían conectarse a Internet de forma que mantuvieran la conectividad en base a una dirección IP matriz (i.e., *Home Address*, HoA) aun cuando utilizaran diferentes redes de acceso, y aun cuando, en el transcurso de una sesión, cambiaran de red de acceso. MIP es una solución de movilidad en la capa de red que cubre tanto el *handover* como la gestión de la localización de los terminales. El planteamiento de MIP se fundamenta en una solución de re-direccionamiento basada en la utilización de una función denominada *Home Agent* (HA) que permite asociar (i.e., *binding*) la dirección *HoA* de un usuario con la dirección IP temporal que tenga asignada ese usuario en la red de acceso que le proporciona conectividad. En MIP, esta dirección provisional recibe el nombre de *Care-of-Address* (CoA). Bajo este planteamiento, el terminal con soporte MIP es el encargado de mantener actualizada la asociación entre HoA y CoA del HA (procedimiento denominado como *binding update*). En este sentido, el protocolo MIP puede enmarcarse dentro de los protocolos de movilidad IP denominados *host-based*, debido al papel activo del terminal en la operación del protocolo. En los últimos años, sobre la base establecida por el protocolo MIP, se han propuesto múltiples complementos y extensiones en aras a mejorar las prestaciones del protocolo MIP tanto en redes IPv4 como en redes IPv6. Así, como ejemplos más significativos, en la RFC 3024 [10] se propone una solución de *reverse tunneling* para MIP sobre redes IPv4 como mecanismo de optimización de rutado. En el caso de redes IPv6, en la RFC 3775 [11] se especifica un nuevo protocolo denominado MIPv6 que ha sido especialmente diseñado para aprovechar las capacidades ofrecidas por la capa de red IPv6. Por otro lado, en aras a facilitar el uso de protocolos de movilidad IP en escenarios donde co-existan redes IPv4 e IPv6, la extensión *Dual Stack MIPv6* (DSMIPv6) del protocolo MIPv6 permite soportar tanto direcciones IPv4 como IPv6. El estándar DSMIPv6 es uno de los soportados en la red troncal EPC en caso de acceder a los servicios mediante redes no 3GPP. El protocolo DSMIPv6 se especifica en RFC 5555 [12]. También destacable es

una extensión de MIPv6 denominada *Hierarchical MIPv6* (HMIPv6) pensada para mejorar las prestaciones del protocolo durante los cambios de CoA en escenarios de movilidad localizada. Una descripción más detallada de los diferentes protocolos de movilidad basados en MIP puede encontrarse en [13].

De forma más reciente se han planteado soluciones de movilidad en la capa de red basadas en MIP donde el terminal no participa de forma directa en la señalización asociada al protocolo MIP. Un claro exponente de este tipo de soluciones, conocidas como soluciones *network-based*, es el protocolo *Proxy MIPv6* (PMIPv6) que ha sido incorporado, entre otros, en redes 3GPP y Mobile WiMAX [14]. En el caso de redes 3GPP, tal como se ha visto en el Capítulo 2, la interfaz entre las pasarelas S-GW y P-GW admite una implementación basada en PMIPv6. Además, tal como se verá más adelante en este mismo capítulo, las soluciones de *interworking* entre LTE y redes no 3GPP también incorporan el uso de PMIPv6 como pieza clave para la provisión de continuidad de servicio entre redes. El protocolo PMIPv6 se especifica en la RFC 5213 [15]. La extensión del protocolo PMIPv4 para soportar tanto direcciones IPv6 como IPv4 (de forma análoga al protocolo DSMIPv6 comentado previamente) no está todavía finalizada en el momento de redacción de este libro [16].

Como alternativa a las soluciones de *handover* basadas en la capa de red, es importante mencionar también que la continuidad del servicio también podría plantearse como una funcionalidad básica integrada en servicios finales. En este sentido, en el contexto de servicios soportados sobre plataformas IMS, el protocolo *Session Initiation Protocol* (SIP), junto con funciones de registro y localización ubicadas en servidores de dicha plataforma, pueden constituir la base de una solución de movilidad en la capa de aplicación. En este tipo de soluciones, si un terminal cambia de dirección IP durante el transcurso de un servicio (e.g., servicio de Voz sobre IP), la propia aplicación es capaz de, mediante la señalización correspondiente, informar del cambio a las entidades participantes en el servicio de forma que los flujos de datos puedan re-establecerse de forma automática (sin intervención del usuario) a través de la nueva dirección IP. En la referencia [17] se analiza el diseño de una solución de movilidad basada en SIP para servicios proporcionados en plataformas IMS. No obstante, las prestaciones en términos de tiempo de interrupción de las soluciones de *handover* basadas en la capa de aplicación suelen ser peores que las soluciones de movilidad sustentadas en la capa de red.

7.2.5.4 Mecanismos de optimización de handover en la capa de red

Estos mecanismos se enmarcan dentro del Nivel D. Los mecanismos de *handover* en la capa de red y de aplicación indicados en el apartado anterior no están diseñados para poder satisfacer los requisitos en caso de querer proporcionar una continuidad transparente o *seamless mobility*. En particular, durante un procedimiento de *handover* deben de efectuarse una serie de pasos asociados a la operativa de la capa de red (e.g., detección del movimiento, adquisición de una nueva dirección y actualización de la asociación entre HoA y CoA) que pueden ocasionar un periodo de interrupción de la conectividad donde el terminal sea inca-

paz de transmitir o recibir información. Este intervalo de interrupción se conoce comúnmente como tiempo de latencia del procedimiento de *handover*. Para minimizar los efectos de dicha transición en la calidad de las conexiones activas, existen varias propuestas destinadas a reducir en la medida de lo posible la latencia del mecanismo de *handover* asociada a la operativa de la capa de red. A modo de ejemplo, en la especificación RFC 4068 [18] se propone una extensión denominada *Fast Handover for MIPv6* (FMIPv6) que permite, entre otras mejoras, que un terminal pueda adquirir una dirección CoA válida para su nueva ubicación mientras sigue conectado a través de la dirección CoA en uso. Además, RFC 4068 define nuevos mensajes de señalización que permiten ejecutar un mecanismo rápido de actualización de la asociación (i.e., *fast binding update*). Análogamente, también se han propuesto extensiones para minimizar la latencia de *handover* para MIPv4 como las contempladas en RFC 4881 [19] y RFC 4988 [20].

Además, en el caso de que el *handover* tenga que realizarse entre redes pertenecientes a diferentes dominios administrativos, esquemas de pre-autenticación como el propuesto en [21] son necesarios para poder reducir los tiempos asociados a los procesos de autenticación y autorización en la nueva red de acceso. Mediante este tipo de esquemas, un terminal podría registrarse en la red de acceso destino a través de la conexión que mantiene con la red de acceso origen. De esta forma, cuando se ejecuta el cambio de red, el reestablecimiento de los servicios a través de la nueva red puede hacerse de forma más rápida ya que el usuario no tiene que esperar a ser autenticado en la nueva red.

La utilización de mecanismos de transferencia de contextos entre redes (e.g., contextos de seguridad, contextos de mecanismos de compresión de cabeceras, perfiles de QoS, capacidades del terminal, etc.) también puede contribuir a reestablecer de forma más rápida los servicios en la nueva red. Mediante la transferencia de contextos entre redes, la red destino puede omitir algunos procedimientos destinados a restituir dichos contextos y que, en algunos casos, comportan un intercambio de señalización adicional entre la red y terminal que retarda el reestablecimiento del servicio. Un ejemplo de protocolo de transferencia de contextos entre redes es el protocolo denominado *Context Transfer Protocol* (CXTP) definido en RFC 4067 [22]. El protocolo CXTP se plantea como mecanismo genérico para la transferencia de diferentes tipos de contextos entre redes, independientemente de la tecnología de acceso radio utilizada. El protocolo CXTP puede utilizarse conjuntamente con protocolos de movilidad IP basados en MIP de forma que, en paralelo a la señalización asociada a los protocolos MIP, puede realizarse la transferencia de contextos pertinente entre las redes.

Finalmente, los mecanismos habilitados para proporcionar información de potenciales redes de acceso candidatas a través de la red de acceso en la que se encuentre conectado el terminal también pueden contribuir a la optimización de los esquemas de *handover* soportados en la capa de red. Estos mecanismos se conocen comúnmente como mecanismos de descubrimiento y selección de red (i.e., *Network Discovery and Selection*). En el documento RFC 5113 [23] se proporciona una visión detallada del alcance de estos mecanismos en el contexto de la problemática de selección de red y/o punto de acceso. A modo de ejemplo,

mediante los mecanismos de *Network Discovery* podría proporcionarse a los terminales conectados en una red de acceso la información necesaria para que pudieran acceder a los servidores de pre-autenticación comentados anteriormente (e.g., direcciones IP de estos servidores).

7.2.5.5 Mecanismos de optimización de handover en las capas radio

La operativa de las capas radio (i.e., capas de enlace y física) durante un proceso de *handover* también puede introducir una serie de retardos que repercuten en la QoS del servicio proporcionado durante la transición entre redes. Así, el establecimiento de la conexión en la nueva red puede conllevar la realización de varios procedimientos (e.g., en redes IEEE 802.11: búsqueda de nuevos puntos de acceso, autenticación y asociación soportados en la capa MAC) que pueden retardar el reestablecimiento del servicio. En este contexto, existen mecanismos de optimización del procedimiento de *handover* asociados a la operativa de las capas radio que permiten evitar o minimizar la latencia del procedimiento de *handover*. Estos mecanismos, junto con los mecanismos de optimización de *handover* en capa de red vistos en el apartado anterior, son importantes para consecución del Nivel D de *interworking*, especialmente en el caso de terminales que no puedan operar más de una interfaz radio de forma simultánea (i.e., terminales *single-radio*). Al mismo tiempo, los mecanismos de optimización de *handover* en las capas radio pueden contribuir a un uso más eficiente de los recursos radio disponibles en las diferentes redes (e.g., mejora de la fase de iniciación y decisión del *handover*). Entre los mecanismos más relevantes a este nivel se encuentran:

- Mecanismos de control de admisión y reserva de recursos entre redes. Mediante estos mecanismos, la red de acceso en la que se encuentra conectado un terminal puede solicitar a la red de acceso destino, de forma previa a la ejecución del *handover*, que realice el control de admisión (si aplica) y la reserva de recursos pertinente. A su vez, la ejecución de estos mecanismos pueden conllevar la pre-configuración de determinados contextos asociados a ciertas funciones de las capas radio (e.g., parámetros de QoS de los servicios portadores activados) en la red candidata, evitando la necesidad de reestablecerlos por completo cuando el terminal haya realizado el cambio de red. Estos mecanismos son claves en la denominada fase de preparación del mecanismo de *handover* (i.e., fase previa a la ejecución del *handover*).
- Mecanismos de provisión de información de las características y configuración de sistemas radio de otras tecnologías en la vecindad (e.g., canales frecuenciales y secuencias de sincronización utilizadas en los puntos de acceso o estaciones base de otras redes). Esta información permite mejorar el proceso de búsqueda de puntos de acceso alternativos así como agilizar el establecimiento de conexiones a través de otras redes. Esta información puede ser clave tanto en la fase de iniciación del *handover* (e.g., la decisión del *handover* tomada en un terminal puede venir condicionada por la información difundida) como en su fase de ejecución (e.g., el terminal puede haber recibido información relativa a la operativa del canal de acceso aleatorio en la nueva red de

forma que se agiliza su tiempo de acceso, ya que no es necesario que el terminal pierda tiempo escuchando los canales de *broadcast* en la nueva red para adquirir dicha información).

- Mecanismos de control de la realización y envío de medidas de diferentes tecnologías radio. La disponibilidad de medidas enviadas desde los terminales que indiquen la presencia y accesibilidad de redes candidatas de otras tecnologías permiten mejorar el proceso de decisión de un *handover*. Para ello, los protocolos de señalización utilizados en las capas radio deben permitir transferir elementos de información con las medidas realizadas en señales de tecnologías radio diferentes (e.g., un terminal LTE puede enviar al eNB mediciones de la señal recibida de una estación base de UMTS). Además del envío de las medidas, puede ser necesario la utilización de mecanismos que controlen o faciliten la realización de las medidas en los terminales. A modo de ejemplo, en caso de terminales *single-radio* con dificultades para realizar mediciones radio en otras frecuencias mientras el terminal se encuentra transmitiendo/recibiendo información, la realización de medidas de otras redes requiere que las capas radio puedan habilitar intervalos específicos donde el terminal pueda realizar las medidas sin penalizar a los servicios en curso (e.g., *compressed mode* en sistemas UMTS).
- Mecanismos de comprobación de la disponibilidad de recursos en las redes vecinas. Los resultados de dichas comprobaciones pueden mejorar la toma de decisiones a la vez que se pueden evitar las situaciones de intentar realizar un *handover* a una red que no tiene recursos para albergar la nueva conexión.

Muchos de los mecanismos citados en este apartado forman parte de las soluciones de *interworking* soportadas entre redes de acceso 3GPP (i.e., E-UTRAN, UTRAN y GERAN). Por el contrario, el desarrollo de este tipo de mecanismos de optimización de *handover* en las capas radio entre redes especificadas por diferentes organismos sigue presentado importantes retos debido a la menor armonización existente en términos de arquitecturas de red y procedimientos. En esta línea, el comité IEEE 802 aprobó a finales de 2008 el estándar IEEE 802.21 [24] que ofrece un marco para el soporte de procedimientos de *handover* entre diferentes tecnologías radio. El estándar IEEE 802.21 se prevé que tenga un papel importante en la integración de sistemas IEEE 802.11 (i.e, Wi-Fi) y 802.16 (i.e., WiMAX) en el contexto de sistemas *ITU-Advanced* [25]. Sin embargo, tal como se verá en los siguientes apartados, las soluciones de *interworking* consideradas actualmente en 3GPP no hacen uso de este estándar.

7.3 Interworking entre LTE y otras redes 3GPP

7.3.1 Marco general

LTE ha sido diseñado para soportar despliegues de red donde co-existan E-UTRAN, UTRAN y/o GERAN. La solución de *interworking* especificada soporta un Nivel D de *interworking* (véase apartado 7.2.4) para la provisión del servicio de conectividad IP a través de los diferentes accesos 3GPP con terminales *single-radio*, teniendo en cuenta, claro está, las diferentes pres-

taciones que pueden conseguirse en cada una de las redes de acceso en cuanto a latencias y tasas de transferencia en bits/s.

La solución de *interworking* entre redes de acceso 3GPP se sustenta fundamentalmente en la introducción de un punto de anclaje en la red troncal EPC que permite la movilidad entre las redes de acceso sin que el servicio de conectividad IP se vea afectado (i.e., la dirección IP asignada al terminal se mantiene aún cuando éste cambia de red de acceso). La interconexión de UTRAN y GERAN a la red troncal EPC se realiza mediante el soporte de funciones del dominio de paquetes GPRS de la red UMTS. La Figura 7.5 ilustra de forma gráfica el planteamiento adoptado.

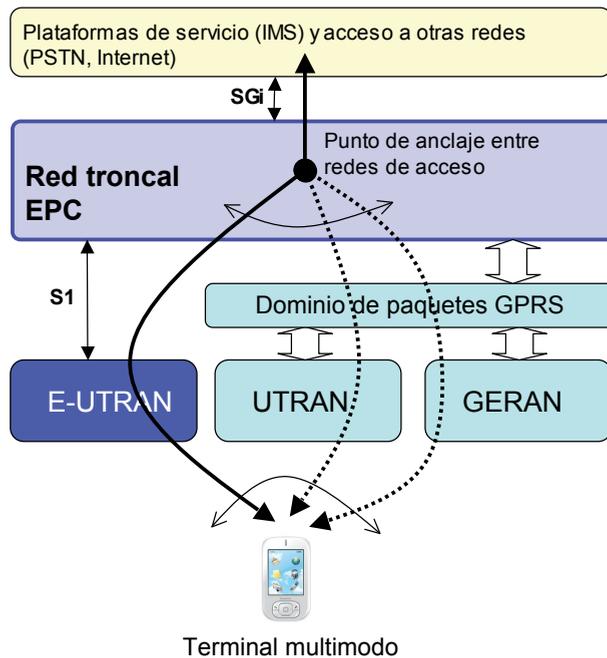


Figura 7.5 Marco de interworking entre redes de acceso 3GPP

Otros aspectos clave de la solución de *interworking* adoptada que se irán desarrollando en posteriores apartados son:

- Las redes de acceso 3GPP tienen mecanismos de difusión de información relativa a la configuración de las estaciones base de otras tecnologías (e.g., a través de los mensajes de difusión de E-UTRAN, se envía información de celdas vecinas de UTRAN y GERAN, y viceversa).
- Se soportan mecanismos de selección de celda (cuando los terminales se encuentran en modos *idle*) que contemplan criterios y parámetros de selección basados en las tres tecnologías [26].

- Se soportan mecanismos de *handover* entre E-UTRAN, UTRAN y GERAN que contemplen el control de admisión y la reserva de recursos en la red de acceso destino de forma previa a la ejecución del *handover*.
- A diferencia de UTRAN y GERAN, no se contempla ninguna interfaz que conecte E-UTRAN con las otras redes de acceso en aras a sustentar el intercambio de información relativa a la gestión de recursos radio entre estas redes (e.g., medidas de carga de las celdas).

Sobre estas premisas, la solución de *interworking* especificada admite dos configuraciones diferentes:

- Utilización de interfaces propias del sistema GPRS. Esta opción se basa en la re-utilización de las interfaces estandarizadas como parte del sistema GPRS para interconectar los equipos de la red troncal GPRS con los equipos de la red troncal EPC. Esta opción hace que sean pocas las modificaciones requeridas en los equipos de las redes GPRS ya desplegadas de cara a su integración en una nueva red LTE. Por el contrario, la utilización de interfaces ya existentes limita la introducción de funciones y procedimientos nuevos que contribuyan a mejorar las prestaciones de la red heterogénea resultante.
- Incorporación de nuevas interfaces entre GPRS y EPC. En esta opción se han definido unas nuevas interfaces que deberían ser soportadas en los equipos de las redes GPRS de cara a su interconexión con una red troncal EPC. La introducción de estas nuevas interfaces conlleva la incorporación de nuevas funcionalidades que permiten, por ejemplo, reducir la carga de señalización de los terminales en modo *idle* (i.e., función denominada como *Idle mode Signalling Reduction*, ISR, que se verá más adelante).

A continuación se describen en detalle las dos configuraciones apuntadas.

7.3.2 Interworking mediante interfaces propias del sistema GPRS

En la Figura 7.6 se muestra la arquitectura de una red troncal EPC a la que se interconectan redes de acceso UTRAN [27] y GERAN [28] mediante el soporte de las interfaces propias del sistema GPRS. A continuación se describen las entidades de red e interfaces del dominio de paquetes GPRS y de las redes de acceso UTRAN y GERAN que forman parte de la solución de *interworking*. Una descripción detallada del resto de entidades de red e interfaces del sistema LTE puede encontrarse en el apartado 2.6 del Capítulo 2.

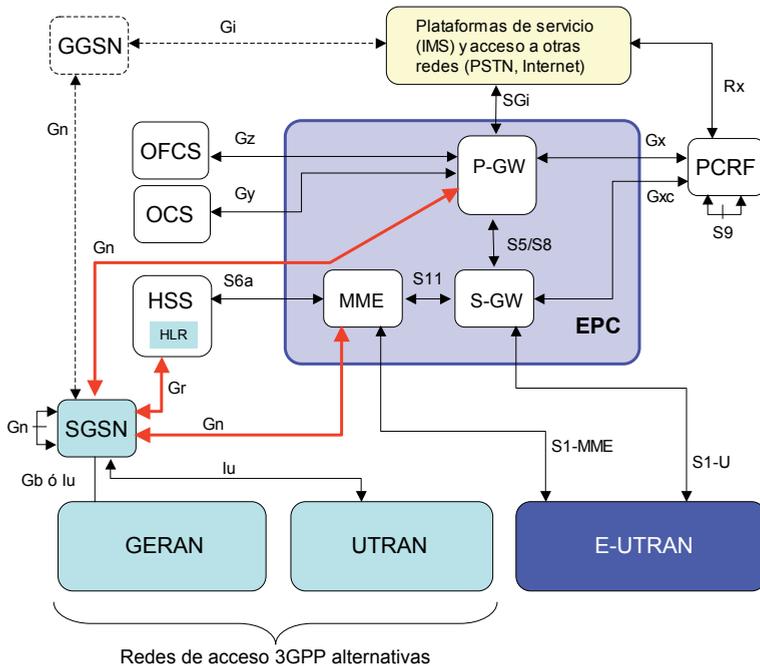


Figura 7.6 Interworking mediante interfaces propios del sistema GPRS

La interconexión entre GERAN y UTRAN y la red troncal EPC se realiza a través de la entidad de red SGSN (*Serving GPRS Support Node*). El nodo SGSN forma parte del dominio de paquetes GPRS [29] y constituye el punto de entrada de la red troncal a los servicios GPRS ofrecidos a través de las redes de acceso UTRAN y GERAN. El nodo SGSN integra tanto las funciones de plano de control como de plano de usuario para la provisión del servicio GPRS. Comparando la arquitectura de la red troncal GPRS con la red troncal EPC, la entidad SGSN de GPRS sería equivalente a las entidades MME y S-GW de EPC. UTRAN se conecta al SGSN mediante una interfaz denominada lu-PS mientras que GERAN puede hacerlo mediante la interfaz Gb (introducida en la primera especificación del sistema GPRS) o bien también a través de una interfaz lu-PS. El nodo SGSN canaliza toda la señalización entre la red GPRS y el equipo de usuario relativa a las funciones de gestión de movilidad y gestión de las sesiones. Para ello, el nodo SGSN accede a la base de datos que contiene la información de los usuarios del sistema, denominada como HLR² (*Home Location Register*) en GPRS, mediante la interfaz Gr. La interfaz Gr es una interfaz basada en el protocolo MAP (*Mobile Application Part*) y que se transporta sobre redes de señalización SS7. Además del nodo SGSN, la implementación del servicio GPRS se sustenta en una entidad de red adicional denominada GGSN (*Gateway GPRS Support Node*). El GGSN es el nodo que interconecta la red troncal GPRS con la red

² En las últimas versiones de los sistemas 3GPP, la funcionalidad HLR se encuentra integrada dentro de la entidad de red denominada como HSS (Home Subscriber Server), tal como se ha visto en el apartado 2.6.2 del Capítulo 2.

externa de paquetes a la que se proporciona acceso (*Packet Data Network, PDN*). Siguiendo la analogía con la red troncal EPC, la entidad GGSN de GPRS es equivalente a la entidad P-GW de EPC. La interconexión entre las entidades SGSN y GGSN se realiza mediante una interfaz denominada Gn a través de la cual se sustentan los procedimientos de gestión de sesiones (e.g., establecimiento de servicios portadores en GPRS, denominados contextos PDP) y movilidad en GPRS (e.g., cambio de SGSN sin perder conectividad con la red externa). La interfaz Gn también se utiliza entre SGSNs para sustentar el cambio de nodo SGSN en una conexión en curso. Así, mediante esta interfaz, un usuario GPRS puede pasar a ser servido por un nuevo SGSN sin que ello afecte a la continuidad del servicio que en todo momento sigue proporcionándose a través del mismo GGSN. Para ello, los nodos SGSNs involucrados se intercambian la información de control necesaria (e.g., contextos asociados al servicio de un usuario) y, de forma opcional, pueden activar un plano de usuario provisional entre ellos para transportar los paquetes de datos del usuario en aras a minimizar la pérdida de información durante el traspaso (funcionalidad equivalente al mecanismo de *Indirect Forwarding Tunnel* utilizado en el mecanismo de *handover* intra-LTE descrito en el apartado 3.3.4 del Capítulo 3). En la parte superior izquierda de la Figura 7.6 se muestra el papel de la entidad GGSN dentro de la arquitectura GPRS, aunque es importante destacar que dicha entidad no forma parte de la arquitectura de *interworking* de UTRAN y GERAN con la red troncal EPC.

Analizados los principales aspectos de los elementos e interfaces propias de GPRS, la solución de *interworking* se articula a través de las interfaces mostradas en rojo en la Figura 7.6. De esta forma, mediante la re-utilización de una interfaz Gn entre SGSN y la pasarela P-GW de la red troncal EPC, la pasarela P-GW es la encargada de realizar la función de punto de anclaje que permite la movilidad entre las redes de acceso 3GPP. Desde la perspectiva del acceso GPRS, la pasarela P-GW se comporta a todos los efectos como un nodo GGSN. Por tanto, a los terminales LTE/UMTS que accedan al servicio de conectividad a través de UTRAN o GERAN, la red GPRS (el nodo SGSN) debe asignarles una pasarela P-GW en lugar de un nodo GGSN convencional. De esta forma, cuando los terminales duales tengan cobertura de la red de acceso E-UTRAN, el servicio de conectividad puede reestablecerse a través de E-UTRAN manteniendo la misma pasarela P-GW para la interconexión con la red externa. El cambio de una red de acceso UTRAN o GERAN a E-UTRAN comporta también el cambio de SGSN por las entidades de la red troncal EPC correspondientes. En este caso, dicho cambio se sustenta en la implementación de la interfaz Gn entre SGSN y MME. La entidad MME alberga el plano de control de la red troncal EPC y a través de ella, puede articularse el cambio hacia o desde la red de acceso E-UTRAN. Nótese que, a diferencia de lo que ocurre en GPRS, en la interfaz Gn entre SGSN y MME no se transportan paquetes de usuario. La interfaz Gn es una interfaz basada en una versión del protocolo GTP (*GPRS Tunneling Protocol*) definida en 3GPP TS 29.060 [46], a diferencia de las interfaces internas del sistema LTE que utilizan una nueva versión del protocolo definida en 3GPP 29.274 [47] (véase el apartado 2.6.3 del Capítulo 2 para una descripción más detallada del protocolo GTP y las interfaces que lo utilizan).

El principio básico de funcionamiento de esta solución es que el servicio de conectividad IP de un usuario del sistema LTE que se conecte a través de las redes de acceso UTRAN o GERAN (debido a, por ejemplo, falta de cobertura de la red E-UTRAN) se establezca a través de la pasarela P-GW apropiada (en lugar de una pasarela GGSN propia de GPRS). Para ello, el SGSN puede hacer uso del conocimiento de las capacidades del terminal en cuanto a qué tecnologías de acceso radio soporta en base a la señalización entre el equipo de usuario y la red. Asimismo, la red GPRS puede seleccionar la pasarela adecuada mediante la utilización del parámetro APN (*Access Point Name*) que forma parte de los datos de subscripción de los usuarios y que también se utiliza en GPRS (en el apartado 3.2 del Capítulo 3 se describe la utilización del parámetro APN en el contexto de LTE). Nótese que el nodo SGSN accede a los datos de subscripción de usuarios LTE a través de la interfaz Gr, tal como se ilustra en la Figura 7.6. Además de los datos de subscripción del usuario, tal como se ha visto en el apartado 2.6.2 del Capítulo 2, en la base de datos HSS se mantiene información necesaria para la operativa del servicio. En este caso, entre los datos almacenados en el HSS (y en particular, la información del HLR) se encuentra la información relativa a la red de acceso radio a través de la cual el usuario se encuentra accesible (e.g., dirección del SGSN para UTRAN y GERAN y de la entidad MME en el caso de E-UTRAN).

7.3.3 Interworking mediante nuevos interfaces entre GPRS y EPC

Esta solución se basa en el soporte de tres nuevas interfaces en la entidad de red SGSN. Estas interfaces son: S3, S4 y S6d. Las interfaces S3 y S4 son interfaces basadas en la misma versión del protocolo GTP que se utiliza en las interfaces internas de la red troncal EPC. En este caso, la interfaz S3 se utiliza para interconectar el SGSN con la entidad MME del plano de control de la red troncal EPC y la interfaz S4 se utiliza entre el SGSN y la pasarela S-GW que constituye el punto de anclaje del plano de usuario dentro de la red EPC. En la misma línea, la interfaz S6d es la adaptación de la interfaz S6a (descrita en el apartado 2.6.2 del Capítulo 2) para el acceso a la base de datos HSS desde el dominio GPRS. El uso de esta interfaz basada en *Diameter* evita la necesidad de soportar la interfaz Gr basada en la señalización MAP/SS7 para el acceso al HLR. Además de estas tres interfaces, la solución planteada permite también la explotación de una interfaz directa entre UTRAN y la red troncal EPC denominada S12. En la Figura 7.7 se muestra la solución de *interworking* especificada, resaltando en color rojo los interfaces a través de los cuáles se articula la provisión y continuidad del servicio LTE a través de las diferentes redes de acceso. A continuación se describen las características más relevantes de esta solución.

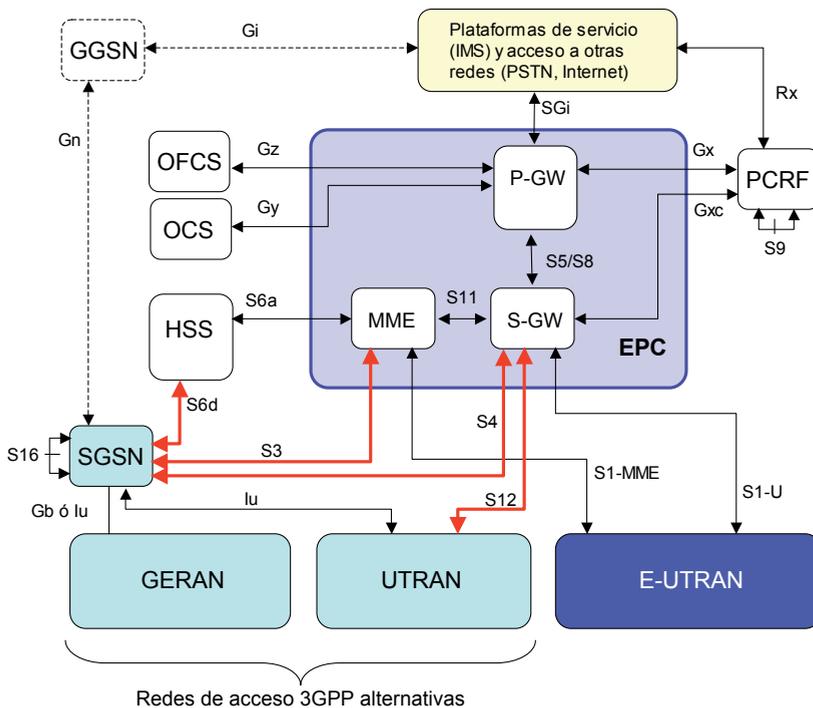


Figura 7.7 Interworking mediante interfaces propios del sistema GPRS

Al igual que en la solución anterior, el punto de anclaje para garantizar la continuidad del servicio de conectividad IP sigue siendo la pasarela P-GW. Sin embargo, en esta solución se fuerza que todo el tráfico de los usuarios atraviese siempre una pasarela S-GW de la red troncal EPC, independientemente de la red de acceso utilizada. Así, para un usuario conectado a través de GERAN o UTRAN, el plano de usuario transcurre a través de un SGSN y un S-GW hasta llegar a la pasarela P-GW. Esta opción es importante en escenarios de itinerancia (véase el apartado 2.6.5 del Capítulo 2) donde la pasarela S-GW de la red visitada siempre constituye un punto de anclaje único de todo el tráfico de usuario. Nótese que en la opción analizada en el apartado anterior, en un escenario de itinerancia, el punto de anclaje en la red visitada sería el SGSN en el caso de acceder a través de UTRAN o GERAN o bien una pasarela S-GW en el caso de E-UTRAN. La utilización de un punto de anclaje único e independiente de la red de acceso en la red visitada facilita al operador de esa red la monitorización y control del tráfico que cursa un usuario en itinerancia. No obstante, la canalización del tráfico a través de una pasarela S-GW de la red troncal conlleva que el tráfico cursado a través de UTRAN y GERAN sea procesado en un elemento de red adicional en comparación con la solución anterior de forma que la latencia de la red puede verse penalizada (en esta solución el plano de usuario atraviesa tres nodos de la red troncal GPRS/EPC mientras que en la anterior solución únicamente eran dos). En el caso de UTRAN, la utilización de la interfaz S12 permite reducir el número de saltos del plano de usuario de forma que se sorte el procesado del plano de

usuario en el SGSN. La interfaz S12 permite el establecimiento del tunel GTP entre el controlador RNC y la pasarela S-GW, evitándose que el tráfico transcurra a través del SGSN. Es importante destacar que, aunque se establezca el túnel del plano de usuario entre UTRAN y S-GW, la señalización de control sigue realizándose a través del SGSN. En el caso de GERAN no se contempla esta opción.

Además de facilitar la operación en escenarios de itinerancia, la utilización de las interfaces S3 y S4 (en lugar de la interfaz Gn) permite incorporar una nueva funcionalidad denominada como *Idle mode Signalling Reduction*, ISR, para gestionar de forma más eficiente la movilidad de los terminales en modo *idle* que pueden estar conmutando entre celdas UTRAN/GERAN y E-UTRAN. La funcionalidad de ISR se detalla a continuación en el apartado 7.3.4.1.

7.3.4 Movilidad entre redes de acceso 3GPP

El marco funcional de gestión de movilidad del sistema LTE especificado en TS 23.401 [30], y analizado en el apartado 3.3 del Capítulo 3 en el contexto de la red de acceso E-UTRAN, cubre también la gestión de movilidad del servicio de conectividad proporcionado por el sistema LTE en escenarios de despliegue en que se utilicen las diferentes redes de acceso 3GPP, atendiendo a las dos soluciones de *interworking* analizadas en las secciones previas.

Tal como ocurre para el caso de E-UTRAN, los mecanismos a través de los cuales se articula la gestión de movilidad entre redes de acceso 3GPP dependen del estado (conectado o “desconectado”) en que se encuentre el terminal. A continuación se describen los mecanismos que se utilizan en cada uno de estos estados.

7.3.4.1 Movilidad de terminales en estado idle

En el sistema LTE, la red troncal EPC es quien gestiona la movilidad de los usuarios que se encuentran en estado *idle*, o equivalentemente, en estado desconectado³. En el contexto de múltiples redes de acceso, un terminal se encuentra en estado *idle* cuando está registrado en la red EPC (y por tanto tiene establecido como mínimo un servicio portador EPS por defecto) pero no se encuentra conectado a través de ninguna red de acceso (i.e., no tiene establecida una conexión RRC con la red de acceso).

En estado *idle*, el terminal es el encargado de seleccionar en todo momento la celda a la que se encuentra sintonizado en aras a recibir la información de sistema (canales de *broadcast*) y avisos de servicio (canales de *paging*). La selección y reelección de celda se lleva a cabo a partir de las medidas realizadas por el terminal junto con los parámetros de selección de celda difundidos por la red. En un escenario con múltiples tecnologías de acceso 3GPP, el terminal realiza medidas de las señales piloto de estaciones base de diferentes tecnologías y

³ Tal como se indica en el apartado 3.3.1 del Capítulo 3, el significado del estado “desconectado” en la gestión de movilidad de la red troncal EPC es equivalente al denominado estado “idle” utilizado en las redes de acceso. En ambos estados, el terminal no dispone de una conexión de control a través de la red de acceso radio. Por este motivo, en este Capítulo se utiliza la notación “idle” como sinónimo del estado “desconectado” de la red troncal EPC.

los parámetros de selección de celda contienen información relativa a las múltiples tecnologías. A modo de ejemplo, en el caso del sistema E-UTRAN, el sistema de información soportado a través de los canales de *broadcast* envía periódicamente mensajes con información sobre frecuencias y celdas vecinas de UTRAN, GERAN y CDMA2000 (i.e., mensajes SIB6, SIB7 y SIB8 descritos en el anexo 2 del Capítulo 5). En la Figura 7.8 se ilustra el mecanismo de (re-)selección de celda de un terminal en estado *idle* en un despliegue multi-RAT.

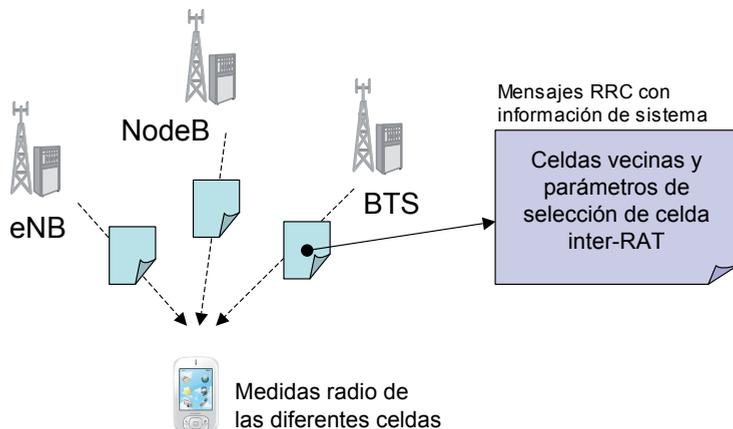


Figura 7.8 Selección de celda de un terminal en estado *idle* en un despliegue multi-RAT

En estado *idle*, el seguimiento de la localización del terminal se realiza mediante la definición de áreas de seguimiento formadas por un conjunto de estaciones base, junto con los mecanismos necesarios para que la red conozca el área o áreas de seguimiento en las que se encuentra accesible el terminal (a través de los canales de *paging*). Tal como se ha visto en el apartado 3.3.2 del Capítulo 3, en el sistema LTE las áreas de seguimiento se denominan *Tracking Areas* (TA) y su actualización la lleva a cabo la entidad MME. En el caso de las redes de acceso UTRAN y GERAN, la gestión de localización se realiza en base a unas áreas de seguimiento denominadas en este caso como *Routing Areas* (RA) y su actualización la lleva a cabo el nodo SGSN. Como resultado del mecanismo de actualización de RA (cuando el terminal sintoniza una celda UTRAN o GERAN) o TA (cuando el terminal sintoniza una celda E-UTRAN), la base de datos del sistema HSS contiene el identificador de la entidad de red SGSN o MME donde se encuentra localizado el usuario. En la Figura 7.9 se ilustra el procedimiento de gestión de localización descrito. Tal como se observa en el gráfico (a) de la Figura 7.9, cuando un usuario sintoniza una celda UTRAN o GERAN, el nodo SGSN dispone de un contexto de movilidad asociado a dicho usuario donde se almacena el identificador de RA. Asimismo, el contexto de movilidad del usuario en la base de datos HSS contiene el identificador del nodo SGSN correspondiente. Además, en el caso de la solución de *interworking* basada en la utilización de las interfaces S3/S4/S6d, la pasarela S-GW también dispone del identificador del nodo SGSN, tal como se ilustra en la Figura 7.9. El motivo es que, en estado *idle*, el plano de usuario únicamente permanece establecido entre la pasarela P-GW y S-GW, de forma que cuando se detecta la presencia de tráfico descendente en la pasarela S-GW, ésta

debe saber a qué nodo notificarlo (e.g., SGSN o MME) para iniciar el procedimiento de aviso. Análogamente, en el gráfico (b) de la Figura 7.9 se muestra la situación en que el usuario se encuentra localizado a través de E-UTRAN.

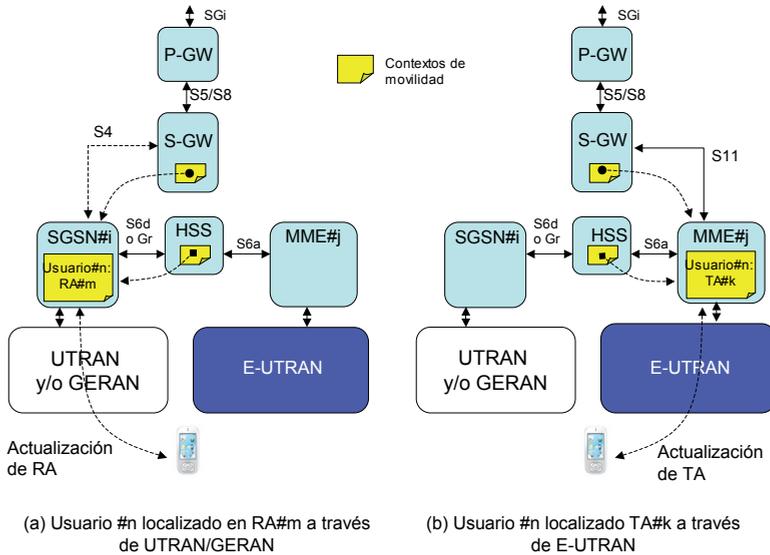


Figura 7.9 Gestión de la localización en un despliegue multi-RAT

Adicionalmente, en el caso de la solución de *interworking* basada en la utilización de las interfaces S3/S4/S6d descrita en el apartado 7.3.3, el sistema LTE soporta la funcionalidad denominada como *Idle Mode Signalling Reduction (ISR)* que permite reducir el número de actualizaciones de RA y TA que se producen en el caso de reelecciones frecuentes entre celdas de E-UTRAN y celdas de UTRAN/GERAN. La idea básica de la funcionalidad ISR es permitir que el terminal pueda mantener dos registros en paralelo, es decir, el terminal se encuentra localizado a la vez mediante un área de rutado y un área (o lista de áreas) de seguimiento. Para ello, es necesario que los nodos SGSN, MME, HSS y S-GW soporten dicha funcionalidad y que las interfaces entre SGSN y las entidades de la red troncal sean S3 y S4 (en lugar de Gn). La activación del modo ISR se produce cuando un terminal, previamente registrado en una RA inicia el registro a una TA, o viceversa. En el procedimiento de actualización de localización el terminal señala que soporta la funcionalidad ISR. Si los elementos de la infraestructura de red citados anteriormente también lo soportan, el procedimiento de actualización termina de forma que tanto SGSN como MME mantienen activo de forma simultánea un contexto de movilidad asociado al usuario (e.g., en el contexto del nodo MME se guarda el TA y el identificador GUTI y en el contexto del SGSN se guarda el RA y P-TMSI). Asimismo, en la base de datos HSS se guardan los identificadores tanto del SGSN como del MME. En cuanto a la entidad S-GW que termina el plano de usuario del terminal en estado *idle*, se establecen ahora dos conexiones de control: una conexión con el MME donde el terminal se encuentra registrado (a través de la interfaz S11) y otra con el SGSN donde también existe un registro activo (a través de la interfaz S4). De esta forma, cuando se recibe tráfico en el S-GW, éste lo notifica tanto

al MME como al SGSN correspondiente, quienes, en paralelo, inician el mecanismo de aviso (i.e., *paging*). Por tanto, el terminal recibirá el mensaje de aviso a través de la red de acceso en la que se encuentre enganchado en ese momento e iniciará un procedimiento de petición de servicio (*Service Request*) que terminará con el reestablecimiento del plano de usuario a través de E-UTRAN, UTRAN o GERAN. Como el reestablecimiento del plano de usuario supone la creación de un túnel GTP entre la pasarela S-GW y la red de acceso correspondiente (e.g., con un eNB en el caso de E-UTRAN, el SGSN en el caso de UTRAN y GERAN o el RNC en el caso de UTRAN con soporte de la interfaz S12), la misma pasarela S-GW se encarga de ordenar la finalización del mecanismo de búsqueda en la entidad MME o SGSN donde no se encontraba accesible el terminal. La configuración de los contextos de movilidad en el caso de un usuario en modo ISR se ilustra en la Figura 7.10.

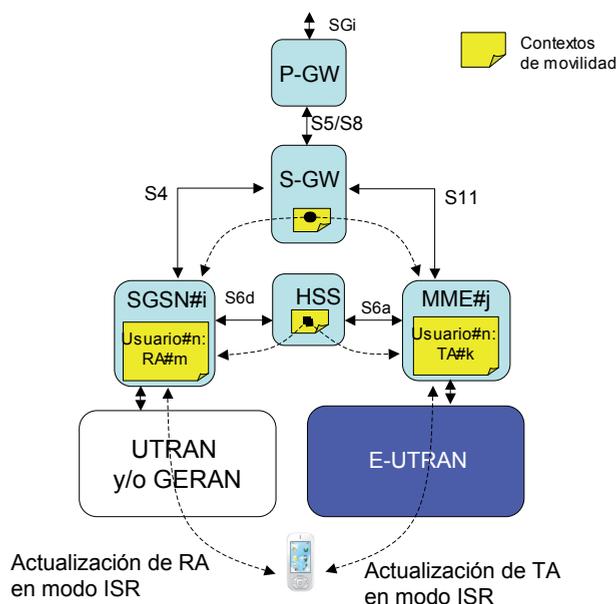


Figura 7.10 Gestión de la localización en un despliegue multi-RAT con el modo ISR

7.3.4.2 Movilidad entre redes de acceso 3GPP en estado activo

Análogamente a la gestión de movilidad en el contexto de una única red de acceso (véase el apartado 3.3 del Capítulo 3), en un escenario multi-RAT, cuando un terminal se encuentra en estado activo, o de forma equivalente, en estado conectado, es decir, tiene establecida una conexión radio con una de las tres posibles redes de acceso 3GPP, la movilidad del terminal entre celdas de diferentes tecnologías se gestiona mediante procedimientos de *handover inter-RAT*.

Los procedimientos de *handover inter-RAT* se sustentan en los mismos principios en los que se basan los procedimientos de *handover* entre celdas de E-UTRAN (véase el apartado 3.3.3 del Capítulo 3 para una descripción detallada). Así, el mecanismo de *handover inter-RAT*

se controla siempre desde la red (i.e., la red de acceso origen decide el cambio hacia una celda de otra red de acceso) teniendo en cuenta medidas enviadas desde los equipos de usuario (i.e., *handover* controlado por la red y asistido por el terminal). De forma previa a la ejecución del *handover*, la red de acceso origen inicia la preparación del *handover* hacia la red de acceso destino a través de la red troncal EPC⁴. La fase de preparación permite reservar recursos en la red de acceso destino (i.e., planteamiento *make before break*) mientras el terminal sigue conectado en la celda de la red de acceso origen. También, la señalización intercambiada en la fase de preparación permite que, en el momento del cambio de celda, el terminal disponga de información específica de la operativa de las capas radio en la celda destino (e.g., parámetros de configuración de los canales de acceso, información de sistema, etc.) que le permita reestablecer la conexión en la nueva red con el mínimo tiempo de interrupción posible. El inicio de la ejecución del cambio de celda en el terminal siempre lo ordena la red de acceso origen.

Los procedimientos de *handover* entre E-UTRAN y UTRAN, y entre E-UTRAN y GERAN, en cualquiera de los sentidos, son bastante similares. En la Figura 7.11 se ilustra un *handover* desde E-UTRAN a UTRAN como un caso representativo de los procedimientos de *handover* inter-RAT. Los diferentes pasos del procedimiento se describen a continuación.

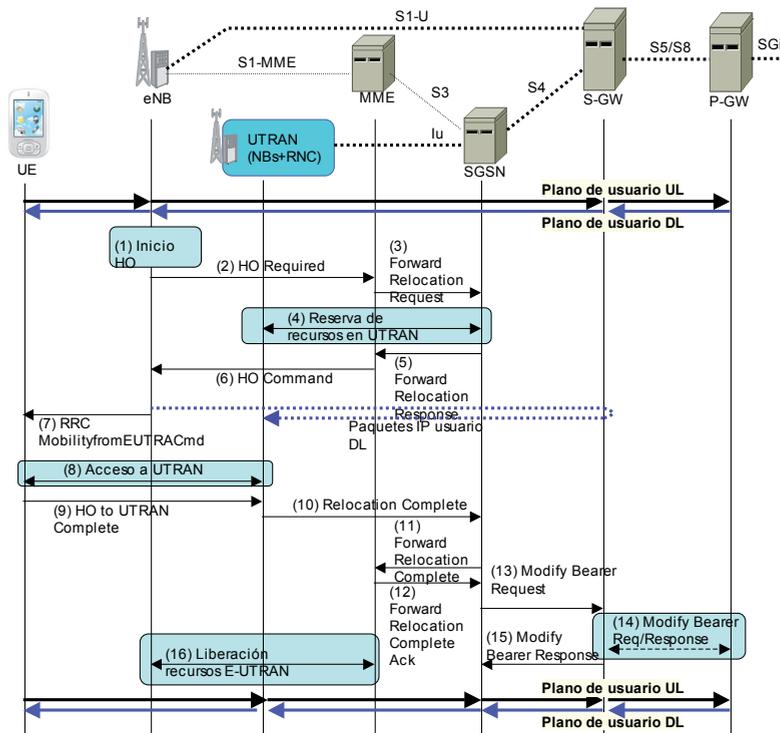


Figura 7.11 Procedimiento de handover desde E-UTRAN a UTRAN

⁴ En handovers entre celdas E-UTRAN, la preparación del handover también puede realizarse directamente entre eNBs en caso de soportarse la interfaz X2. En caso contrario, la preparación del handover también se realiza a través de la red troncal EPC como en el caso de handovers inter-RAT.

(1) Iniciación del *handover*

El procedimiento de *handover* lo inicia el eNB a través del cual se conecta el terminal. La decisión de realizar el *handover* puede ser debida a que, por ejemplo, las medidas recibidas desde el terminal indican la recepción de señal en el terminal de una celda UTRAN con mejores condiciones que la celda actual en E-UTRAN.

(2)-(6) Preparación del *handover*

El eNB origen envía la petición de *handover* (mensaje "*HO required*") hacia la entidad MME de la red troncal. El mensaje contiene información que permite identificar la red de acceso UTRAN (e.g., identificador del equipo RNC de UTRAN) donde se encuentra la celda hacia la que se pretende dirigir el *handover*. A partir de esta información, la entidad MME determina que se trata de un *handover* inter-RAT entre E-UTRAN y UTRAN e inicia el procedimiento de *Forward Relocation* con la entidad SGSN correspondiente. El procedimiento de *Forward Relocation* permite transferir el contexto del usuario (e.g., conexiones PDN, servicios portadores establecidos, contexto de movilidad, etc.) de la entidad MME origen hacia el nodo SGSN destino. De esta forma, el nodo SGSN destino conoce las características de los servicios portadores activos (e.g., parámetros de QoS) y puede proceder a reservar los recursos necesarios en la red de acceso destino UTRAN (i.e., activación de *Radio Access Bearers*, RABs, en UTRAN). El nodo SGSN es el encargado de mapear los parámetros de QoS utilizados en los servicios portadores EPS con los parámetros de QoS utilizados en UMTS. El procedimiento de reserva de recursos en UTRAN (paso (4) en la Figura 7.11) se soporta mediante los procedimientos existentes en la interfaz lu entre SGSNs y RNCs. De hecho, el procedimiento completo de *Forward Relocation* es una extensión del procedimiento ya soportado en UMTS para realizar *handovers* entre SGSNs. Una vez finalizada la reserva de recursos en UTRAN, el nodo RNC destino queda ya preparado para la recepción de paquetes IP del usuario en la dirección descendente.

Finalmente, el SGSN destino responde a la petición de *Forward Relocation* a la entidad MME origen y ésta a su vez señala al eNB la finalización de la etapa de preparación del *handover*. Los mensajes (5) y (6) pueden transportar de forma transparente información relativa a parámetros de configuración de la nueva conexión radio del terminal con UTRAN. En este momento, se pueden establecer túneles en la red de forma que el tráfico IP en sentido descendente que llegue al eNB origen sea redirigido hasta el RNC destino.

(7)-(12) Ejecución del HO en la interfaz radio

El terminal recibe la orden de conmutar su conexión radio hacia una celda UTRAN mediante la señalización de control RRC (e.g., mensaje *MobilityFromEUTRACmd* ilustrado en el paso (7)). A través de esta señalización, el terminal puede recibir los parámetros de configuración radio necesarios para proceder al acceso a través de UTRAN. A partir de este instante, el terminal sincroniza su interfaz radio con la celda correspondiente de UTRAN e inicia el acceso a la red destino. Una vez el terminal ha establecido la nueva conexión en UTRAN, el controlador RNC indica la finalización del proceso de *Relocation* mediante el mensaje "*Relocation*

Complete". En este momento, la entidad SGSN destino puede ya recibir tráfico en dirección ascendente proveniente de UTRAN que se encamina hacia la pasarela S-GW. Finalmente, el nodo SGSN informa a la entidad MME origen del resultado del *handover* para que puede proceder a la liberación de recursos en la red de acceso E-UTRAN origen.

(13)-(15) Actualización de las conexiones en la red troncal EPC

La entidad SGSN interactúa con la pasarela S-GW para actualizar el contexto de datos asociado al usuario y conmutar el túnel GTP todavía establecido en la interfaz S1 con el eNB origen hacia la entidad SGSN a través de la interfaz S4. En caso de que UTRAN soportara la funcionalidad denominada como *Direct Mode*, el túnel GTP podría también establecerse entre la pasarela S-GW y el RNC de UTRAN, evitándose que el tráfico de usuario atravesara la entidad SGSN.

En el ejemplo mostrado en la Figura 7.11 también se ilustra la posibilidad de que la pasarela S-GW informe a la pasarela P-GW del cambio realizado (e.g., notificación de cambio de tecnología de acceso). Con esta información, la pasarela P-GW podría interactuar con el subsistema PCC para, por ejemplo, modificar las condiciones de tarificación o modificar las características de QoS de alguno de los servicios portadores.

(16) Liberación de recursos

Finalmente, la entidad MME indica al eNB origen que el HO se ha realizado con éxito y que puede proceder a la eliminación del contexto y recursos asociados al usuario.

El procedimiento mostrado en la Figura 7.11 corresponde a una solución de *interworking* basada en la utilización de las interfaces S3/S4/S6d en GPRS (descrita en el apartado 7.3.3). En el caso de la solución de *interworking* basada en la reutilización de las interfaces del sistema GPRS (descrita en el apartado 7.3.2), el procedimiento de *handover* también mantiene las mismas pautas. En particular, en este caso, el procedimiento de *handover* es equivalente al procedimiento denominado *Inter SGSN handover* especificado para el dominio GPRS [29] donde la entidad MME de la red troncal EPC realiza las funciones correspondientes al plano de control de un nodo SGSN. Una descripción detallada de todos los procedimientos de *handover* entre redes de acceso 3GPP se proporciona en TS 23.401 [30].

7.3.5 Soluciones de interworking para los servicios de voz

El servicio de voz en las redes 3GPP se soporta mediante dos soluciones completamente diferentes: una basada en la utilización del dominio de conmutación de circuitos de la red troncal GSM/UMTS y la otra soportada en el subsistema IMS, accesible a través del dominio de conmutación de paquetes (e.g., GPRS y EPC). En el caso de las redes de acceso UTRAN y GERAN, ambas redes proporcionan acceso tanto al dominio de circuitos como al dominio de paquetes de forma que la provisión del servicio de voz admite ambas soluciones. Por el contrario, la red de acceso E-UTRAN no proporciona acceso al dominio de circuitos con lo que la solución IMS constituye la elección natural para soportar el servicio de voz. El servicio de voz que ha sido estandarizado sobre IMS se conoce como *Multimedia Telephony* (MMTel) [31]. El

servicio MMTel, además de la transferencia de la voz, introduce posibilidades adicionales tales como el envío de información de texto o imágenes junto con la voz.

En cualquier caso, en las fases iniciales de despliegue de una red E-UTRAN, es bastante probable que un operador no pueda proporcionar una cobertura completa y contigua del área de servicio. Por ello, de cara a la provisión de un servicio de voz con continuidad de servicio en toda el área de servicio, el servicio de voz MMTel ofrecido a través de la red E-UTRAN debe complementarse con el servicio de voz del dominio de circuitos en las zonas donde E-UTRAN no tenga cobertura.

El mecanismo denominado como *Voice Call Continuity* (VCC) permite proporcionar continuidad del servicio de voz entre el dominio de circuitos de GSM/UMTS y el servicio MMTel soportado en IMS. VCC se ha desarrollado como una aplicación de IMS que se ejecuta en la plataforma IMS del operador matriz. A través de dicha aplicación, las llamadas de un terminal que soporte VCC, denominado simplemente como terminal VCC, pueden ser ancladas en la plataforma IMS del operador matriz, tanto si las llamadas se inician desde el dominio de circuitos como a través de IMS. El anclaje se consigue mediante la utilización de mecanismos de re-encaminamiento de llamadas que ya existen en el dominio de circuitos. De esta forma, una llamada originada por un terminal VCC en el dominio de circuitos se encamina hacia la plataforma IMS de su operador, y a partir de aquí, se continúa el encaminamiento de la llamada hasta el usuario destinatario. Así pues, gracias a este anclaje resulta posible que una llamada iniciada por un terminal VCC a través del dominio de circuitos (e.g., llamada de voz iniciada en UTRAN) sea transferida hacia el dominio de paquetes a IMS sin interrumpir el servicio (e.g., continuación de la llamada en E-UTRAN) y viceversa. Las transferencias entre dominios (CS y IMS) pueden realizarse en las dos direcciones y, en el transcurso de una sesión, la aplicación VCC puede soportar múltiples cambios. A modo ilustrativo, en el gráfico (a) de la Figura 7.12 se muestra el modelo de *interworking* de VCC donde se muestra como una llamada establecida a través del dominio de circuitos de GSM/UMTS, en lugar de ser encaminada a través de la red de circuitos (vía la pasarela *Gateway MSC*, GMSC), se encamina hacia la plataforma IMS. Allí, la aplicación VCC realiza funciones de anclaje de la señalización asociada a la llamada. De esta forma, si durante el transcurso de la llamada el terminal cambia su conexión hacia una red de acceso E-UTRAN, el servicio de voz con la red telefónica puede mantenerse a través del subsistema IMS, tal como se ilustra en el gráfico (b) de la Figura 7.12. Una descripción más detallada de los elementos genéricos de la plataforma IMS y sus principales funciones puede encontrarse en el apartado 2.7 del Capítulo 2. Asimismo, la descripción detallada de la arquitectura funcional de VCC y los procedimientos asociados se proporciona en el documento 3GPP TS 23.206 [32]. Es importante destacar que la función de VCC actualmente sólo está definida para llamadas de voz, y por tanto no aplica a otros servicios proporcionados sobre el dominio de circuitos como, e.g., videollamadas.

La solución VCC no es propia del sistema LTE sino que ya se introdujo para resolver la misma problemática de continuidad de llamadas entre el dominio de circuitos y el subsistema IMS en el contexto de redes UMTS. Una limitación importante de la solución VCC es que está

planteada para ser utilizada con terminales *multimodo dual-radio*. El hecho de utilizar terminales *dual-radio* hace que no se requiera una interacción entre las redes troncales de conmutación de circuitos y de conmutación de paquetes (GPRS o EPC) de cara a soportar el cambio de red de acceso. Sin embargo, tal como sucede con los actuales terminales multimodo 2G/3G, cabe esperar que la mayoría de terminales multimodo que incorporen soporte para E-UTRAN sean también terminales *single-radio*. Por este motivo, como mejora de la solución VCC, recientemente 3GPP ha especificado la solución conocida como *Single Radio VCC* (SRVCC) [33]. La solución SRVCC se fundamenta en el mismo principio de funcionamiento de la solución VCC descrito en este apartado con la salvedad de que la solución puede ser aplicada en el caso de terminales *single-radio*. Para ello es necesaria la introducción de nuevas interfaces de red que interconecten la entidad de red MME de la red troncal EPC con las entidades de conmutación MSC del dominio de circuitos. Una explicación más detallada de las soluciones VCC y SRVCC, así como otras propuestas que se han planteado para el soporte del servicio de voz sobre LTE (e.g., *Circuit-Switched Fallback*, CSF) quedan fuera de ámbito de este libro. El lector interesado en este tópic puede encontrar una buena referencia en [34].

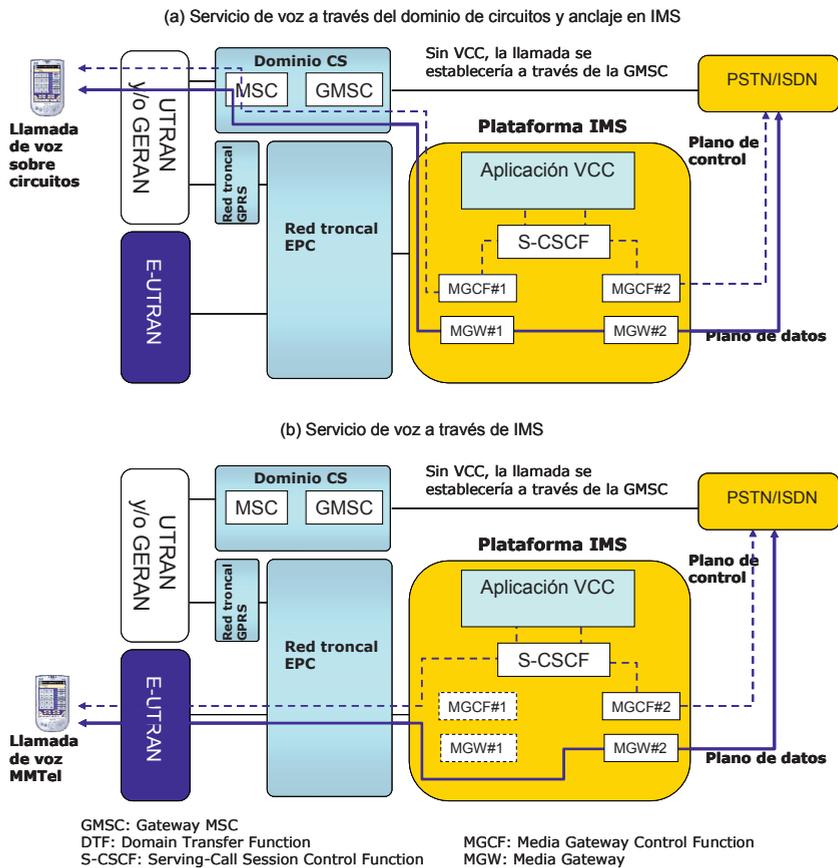


Figura 7.12 Modelo de interworking en VCC

7.4 Interworking entre LTE y otras redes no 3GPP

7.4.1 Marco general

La red troncal EPC ha sido diseñada para poder proporcionar sus servicios de conectividad IP a terminales que utilicen redes de acceso radio o cableadas diferentes a las especificadas por el 3GPP [35]. En la Figura 7.13 se ilustra el marco de *interworking* del sistema LTE con otras redes de acceso no 3GPP. La red troncal EPC, junto con otros elementos comunes estandarizados por el 3GPP (base de datos HSS, subsistema PCC, subsistemas de facturación, etc.), constituyen el núcleo de la arquitectura de red heterogénea donde se sustentan las funciones transversales necesarias para la provisión del servicio y que son independientes de la red o redes de acceso utilizadas. En particular, a través de la infraestructura 3GPP, se sustentan las funciones relacionadas con la gestión de usuarios (e.g., almacenamiento y acceso a los datos de suscripción, control de acceso, etc.), el control de QoS y tarificación del servicio de conectividad, y el soporte de movilidad entre las diferentes redes de acceso. El marco genérico desarrollado no impone ninguna restricción a las redes de acceso no 3GPP consideradas, salvo el hecho de que sean redes que proporcionen un mecanismo de transferencia de paquetes IP. Ejemplos de posibles redes de acceso son: redes de acceso radio CDMA2000, redes de banda ancha cableadas con terminaciones inalámbricas (e.g., servicio ADSL con puntos de acceso inalámbricos IEEE 802.11) y redes de acceso Mobile WiMAX. El marco planteado no presupone ninguna interacción entre las diferentes redes de acceso, es decir, las redes de acceso pueden estar completamente desacopladas unas de otras. Por tanto, la red troncal EPC actúa como punto de anclaje del servicio de conectividad IP, ofrecido a través de cualquiera de las redes de acceso 3GPP y no 3GPP, garantizándose la continuidad del servicio al cambiar de red de acceso (e.g., la dirección IP asignada al terminal se mantiene aún cuando éste cambia de red de acceso).

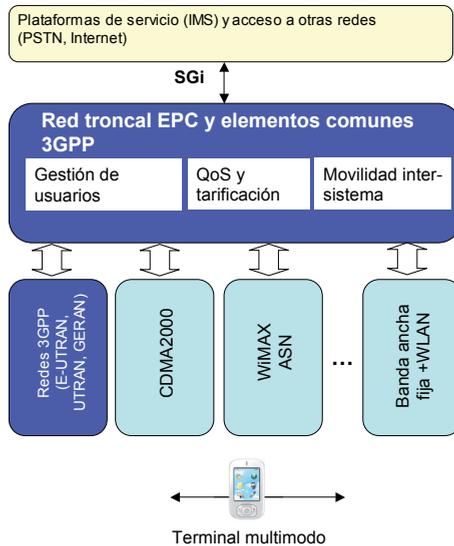


Figura 7.13 Marco de interworking con redes de acceso no 3GPP

Atendiendo a la diversidad de las posibles redes de acceso que pueden utilizarse, así como a las diferentes opciones tecnológicas existentes para sustentar la movilidad en la capa de red entre sistemas (véase el apartado 7.2.5.3), 3GPP ha optado por especificar varias soluciones de *interworking* entre el sistema LTE y las redes no 3GPP. Las posibles soluciones pueden clasificarse en torno a dos aspectos fundamentales:

- La red de acceso no 3GPP se considera *trusted* o *untrusted*.
- El protocolo de movilidad de capa de red utilizado *network-based* o *host-based*.

A continuación se describen ambos aspectos.

El término *trusted* se utiliza para indicar que la red de acceso no 3GPP dispone de los mecanismos de seguridad apropiados que hacen que el operador de la red troncal EPC a la que se conecta la considere como tal (e.g., mecanismos de autenticación del usuario, servicios de confidencialidad e integridad, etc.)⁵. Ejemplos de redes *trusted* podrían ser redes de acceso CDMA2000 o Mobile WiMAX conectadas a la infraestructura de la red troncal EPC mediante una red segura. En la Figura 7.14 se ilustra un escenario de *interworking* con una red de acceso *Mobile WiMAX* (denominada *Access Service Network* según las especificaciones *WiMAX NWG Release 1* [36]). La Figura 7.14 muestra como la información cursada por un terminal conectado a través de la red de acceso *Mobile WiMAX* (utilizando las credenciales de su USIM) es encaminada desde la pasarela *ASN GW (ASN Gateway)* de la red de acceso hasta la red troncal EPC desde donde se accede a las redes externas (e.g., Internet, plataformas de servicios IMS). Los detalles de este tipo de solución se abordan en el apartado 7.4.2.

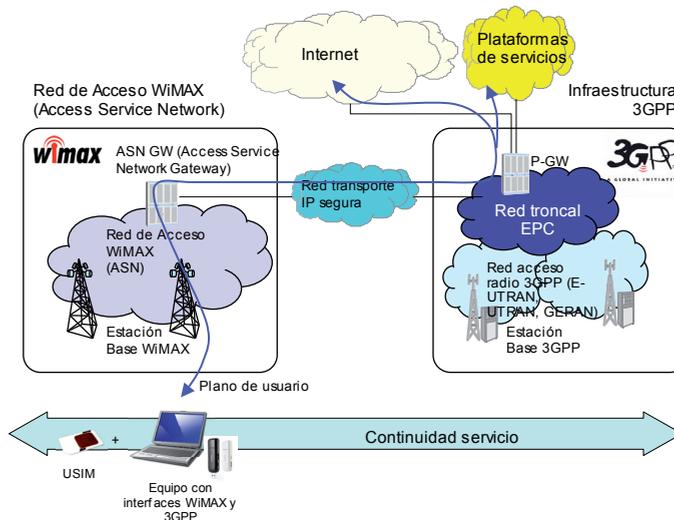


Figura 7.14 Ilustración de un escenario de interworking con una red de acceso *Mobile WiMAX*

⁵ Las especificaciones 3GPP no definen los requerimientos técnicos que debe cumplir una red no 3GPP para ser considerada como *trusted*. Así, la consideración como tal está sujeta al criterio del operador que ofrece su servicio de conectividad IP a través de dicha red.

Por el contrario, una red *untrusted* es aquella en la que el operador de la red troncal no tiene garantías suficientes para considerarla como *trusted* y por tanto requiere de la introducción de medidas adicionales para garantizar el acceso al servicio proporcionado y evitar posibles vulnerabilidades o usos no autorizados. Un ejemplo de red *untrusted* podría ser una red WLAN pública o de uso privado en una oficina u hogar que, a través de Internet, se utilizara para acceder al servicio de conectividad IP proporcionado por la red troncal EPC de un operador LTE. Una ilustración de un escenario de *interworking* con una red Wi-Fi de acceso público se muestra en la Figura 7.14. Como diferencias principales entre este escenario y el correspondiente a la red de acceso Mobile WiMAX visto en la Figura 7.14 cabe destacar que la conexión entre la red de acceso Wi-Fi y la infraestructura de la red 3GPP se realiza ahora a través de una red pública como es Internet. Además, cabría considerar la posibilidad de que la red Wi-Fi no tuviera los mecanismos de seguridad habilitados (e.g., servicio Wi-Fi en modo abierto habilitado a los clientes de cafetería). En este caso, tal como se verá en los detalles de esta solución abordada en el apartado 7.4.3, la infraestructura de la red troncal EPC incorpora unas pasarelas específicas con soporte de protocolos IPsec para permitir un acceso al servicio de conectividad IP con las garantías de seguridad necesarias.

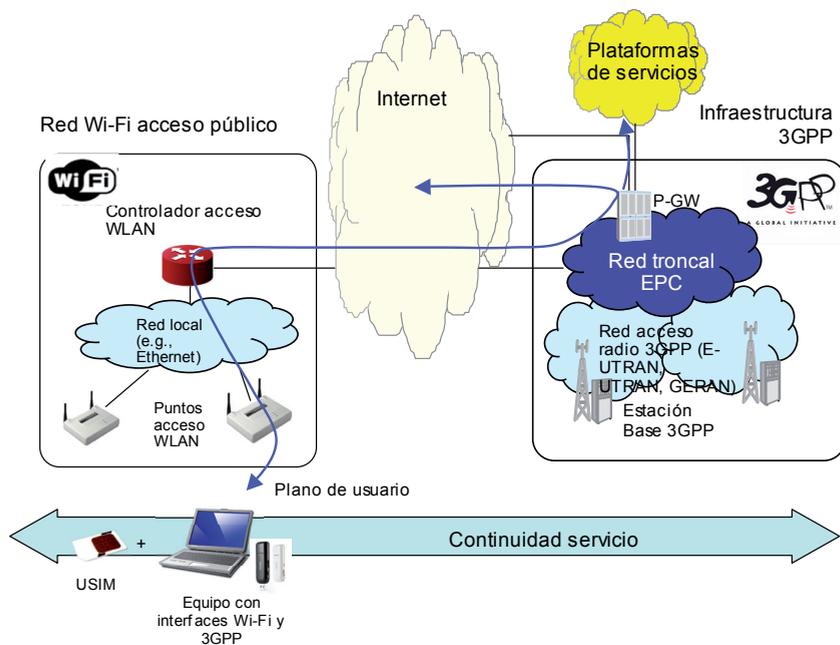


Figura 7.15 Ilustración de un escenario de interworking con una red Wi-Fi de acceso público

En cuanto a la movilidad *network-based* o *host-based*, tal como se ha visto en el apartado 7.2.5.3, se denomina movilidad *network-based* aquella solución de movilidad donde el terminal no participa de forma directa en la señalización asociada a la gestión de movilidad.

Ejemplos de protocolos que proporcionan movilidad *network-based* son Proxy MIPv6 y GTP⁶. Por otro lado, la movilidad *host-based* es aquella en la que existe un cliente de movilidad instalado en el terminal y a través del cual se controla la configuración de la interfaz IP en el cambio de red de acceso. Los protocolos MIPv4 y MIPv6 son dos esquemas básicos que proporcionan movilidad *host-based*. Las ventajas principales de los protocolos de movilidad *network-based* frente a los de movilidad *host-based* es que los primeros requieren de terminales menos complejos (sin cliente de movilidad) y, al evitarse la transferencia de señalización del protocolo a través de la interfaz radio, las prestaciones en términos de latencia en el proceso de actualización de la asociación (i.e., *binding update*) son mejores. En cambio, la utilización de estas soluciones requiere que la red de acceso proporcione soporte para ello. Por otro lado, las soluciones de movilidad en el terminal pueden utilizarse con independencia del soporte de movilidad que proporcione la red de acceso. En particular, los protocolos de movilidad especificados por 3GPP para el acceso desde redes no 3GPP son: DSMIPv6, PMIPv6, MIPv4.

En la Tabla 7.1 se indican las entidades de red e interfaces adicionales que incorpora la red troncal EPC para sustentar las diferentes opciones de *interworking* en función de los dos aspectos descritos anteriormente (red no 3GPP *trusted* o *untrusted* y movilidad en capa de red *network-based* o *host-based*). Los detalles de las diferentes opciones se analizan en los siguientes apartados. En particular, en el apartado 7.4.2 se analiza la solución genérica para redes *trusted* para las dos posibles opciones de movilidad y en el apartado 7.4.3 se cubre la solución para redes *untrusted*, donde también se consideran las dos posibles opciones de movilidad. Las dos soluciones genéricas están diseñadas para proporcionar un Nivel C de *interworking* (véase el apartado 7.2.4). Posteriormente en el apartado 7.4.4 se describen aspectos básicos de la gestión de movilidad que son comunes a ambas soluciones genéricas y, por último, en el apartado 7.4.5 se analizan los detalles más relevantes de una solución específica de *interworking* optimizada para el caso de redes CDMA2000. Tal como se verá, la solución optimizada es una extensión de la solución genérica para redes *trusted* que permite conseguir un Nivel D de *interworking* para terminales *single-radio* entre redes 3GPP y CDMA2000.

⁶ Tal como se ha visto en el apartado 2.6.3 del Capítulo 3, además de soportar movilidad entre las pasarelas de la red troncal EPC, el protocolo GTP también soporta mecanismos de gestión de las sesiones y servicios portadores.

Tabla 7.1 Entidades de red e interfaces adicionales para las soluciones de interworking entre LTE y redes de acceso no 3GPP

Opciones de interworking		Trusted		Untrusted	
		Movilidad network-based (PMIPv6)	Movilidad host-based (DSMIPv6)	Movilidad network-based (PMIPv6, MIPv4)	Movilidad host-based (DSMIPv6)
Entidades de red adicionales		Servidor AAA (TS 23.402 [35], TS 23.002 [37])		Servidor AAA ePDG (TS 23.402 [35], TS 23.002 [37])	
Interfaces adicionales	Acceso a la base de datos HSS	SWx (TS 29.273 [38])			
	Soporte movilidad P-GW (e.g., actualización de la localización)	S6b (TS 29.273 [38])			
	Control de acceso	STa (TS 29.273 [38])		SWa (TS 29.273 [38])	
	Control de QoS y tarificación	Gxa (TS 23.203 [39])		Gxb (no especificado todavía)	
	Movilidad P-GW y red de acceso o terminal	S2a (TS 29.275 [40])	S2c (TS 24.303 [41])	S2b (TS 29.275 [40] y TS 24.304 [42])	S2c (TS 24.303 [41])
	Servicios de seguridad complementarios	Ninguno		SWu, SWn, SWm (TS 29.273 [38])	

Las referencias a documentos 3GPP proporcionadas en la tabla son las más representativas aunque no constituyen una lista exhaustiva de todos los documentos 3GPP relacionados con cada aspecto.

7.4.2 Solución genérica para redes Trusted

La arquitectura de red de la solución de *interworking* genérica para redes no 3GPP *trusted* se muestra en la Figura 7.16. A continuación se describen las entidades de red e interfaces adicionales introducidas en esta solución. La descripción detallada del resto de entidades de red e interfaces del sistema LTE puede encontrarse en el apartado 2.6 del Capítulo 2.

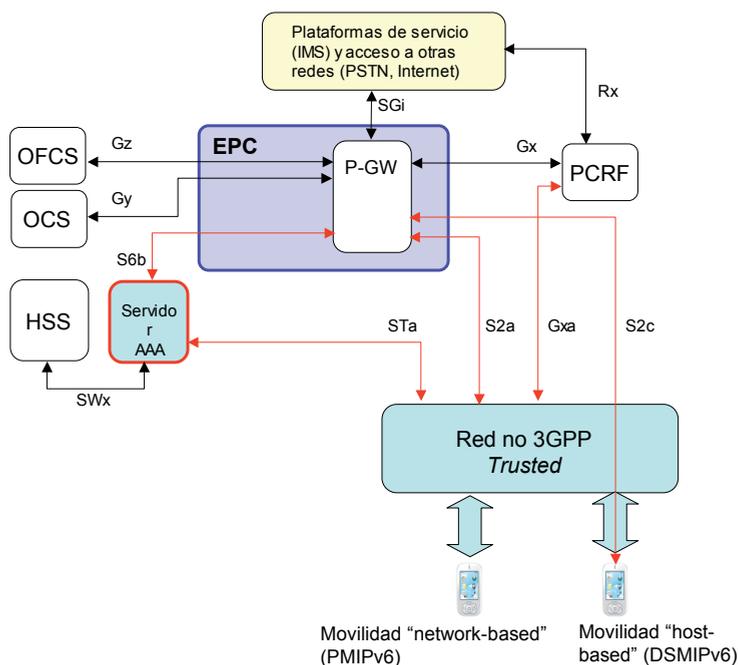


Figura 7.16 Solución genérica para redes no 3GPP *trusted*

Los datos de suscripción almacenados en la base de datos HSS (i.e., vectores de autenticación, identificadores APN, parámetros de QoS y tarificación, direcciones IP asignadas, etc.) se utilizan para gestionar el acceso de los usuarios a través de red no 3GPP. Para ello, se define una nueva entidad de red denominada *Servidor AAA* que interactúa con la base de datos HSS a través de una interfaz denominada SWx basada en el protocolo *Diameter*⁷. Mediante la interfaz SWx, el servidor AAA puede acceder a los datos del perfil de suscripción de un usuario. El servidor AAA se conecta con la red de acceso no 3GPP mediante la interfaz STa. Esta interfaz se utiliza para sustentar los procedimientos de control de acceso (e.g., autenticación y autorización de los usuarios en la red no 3GPP) así como la configuración de los protocolos de movilidad (e.g., transferencia de parámetros tales como la dirección del *Home Agent* para la movilidad basada en DSMIPv6). La interfaz STa está basada también en el protocolo IETF Diameter.

Además de dar soporte a las funciones de autenticación y autorización en la red de acceso, el servidor AAA también proporciona acceso a datos relativos a la suscripción de los usuarios a la pasarela P-GW. Estos datos de suscripción son utilizados por la pasarela P-GW para gestionar el servicio de conectividad IP (e.g., perfil de QoS suscrito para redes no 3GPP). La comunicación entre ambas entidades de red se realiza mediante la interfaz S6b basada en *Diameter*. A través de esta interfaz, la pasarela P-GW también indica al servidor AAA su iden-

⁷ Véase el apartado 2.6.3.3 del Capítulo 2 para una descripción más detallada del protocolo Diameter y su aplicación en diferentes interfaces de la red LTE.

tividad (e.g., dirección IP de la pasarela) y la identidad de la red externa APN a la que se haya conectado un usuario. Esta información el servidor AAA la almacena en el HSS para sustentar el mecanismo de *handover* entre redes de acceso (el HSS guarda la dirección de la pasarela y la red externa a la que el usuario está conectado), tal como se verá más adelante en el apartado 7.4.4. Además, a través de esta interfaz, la pasarela P-GW lleva a cabo procedimientos relacionados con la operativa de los protocolos de movilidad que terminan en la pasarela (e.g., autenticación y autorización del usuario necesaria durante el procedimiento de registro de los diferentes protocolos MIP soportados).

La movilidad entre redes de acceso se fundamenta en el mantenimiento de la pasarela P-GW como punto de anclaje del servicio de conectividad, junto con la utilización de un protocolo de movilidad en la capa de red a través de la red no 3GPP. En particular, la solución genérica para el acceso *trusted* a redes no 3GPP admite las siguientes opciones:

- Interfaz S2a para movilidad *network-based*. Los protocolos soportados en esta interfaz son PMIPv6⁸ y MIPv4⁹ (modo *Foreign Agent*). En la red troncal EPC la interfaz termina en la pasarela P-GW (que sustenta las funciones de *Local Mobility Anchor*, LMA, en el caso de PMIPv6 y de *Home Agent*, HA, en el caso de MIPv4). En la red de acceso, la interfaz termina en la pasarela que albergue las funciones correspondientes asociadas a dichos protocolos (funciones de *Mobile Access Gateway*, MAG, en el caso de PMIPv6, y de *Foreign Agent*, FA, en el caso de MIPv4). En la Figura 7.17 se ilustra la torre de protocolos para los planos de control y usuario correspondientes al uso de esta interfaz en el caso del protocolo de movilidad PMIPv6. Nótese como el terminal no participa en el plano de control sino que es la red de acceso no 3GPP quien se ocupa de gestionar la operativa del protocolo PMIPv6. Asimismo, el túnel del protocolo de movilidad en el plano de usuario se establece desde la red de acceso no 3GPP hasta la pasarela P-GW (el mecanismo de túnel utilizado en PMIPv6 es *Generic Routing Encapsulation*, GRE).
- Interfaz S2c para movilidad *host-based*. El protocolo soportado en esta interfaz es DSMIPv6, junto con IPsec y IKEv2¹⁰ (*Internet Key Exchange version 2*) para proporcionar servicios de seguridad a la señalización del protocolo de movilidad. La interfaz se define entre la pasarela P-GW de la red troncal EPC y el propio equipo de usuario. La pasarela P-GW alberga las funciones de HA del protocolo DSMIPv6 y en el terminal se aloja el cliente DSMIPv6. La torre de protocolos para los planos de control y usuario correspondientes al uso de esta interfaz se muestran en la Figura 7.18. A diferencia de la torre de protocolos de la interfaz S2a, el terminal alberga las funcionalidades de un cliente DSMIPv6 para gestionar la operativa del protocolo de movilidad. Nótese también como ahora el túnel del protocolo de movilidad en el plano de usuario se extiende desde el terminal hasta la pasarela P-GW de la red troncal.

⁸ Véase el apartado 2.6.3.4 del Capítulo 2 para detalles adicionales sobre el protocolo PMIPv6

⁹ Véase el apartado 7.2.5.3 de este Capítulo para detalles adicionales sobre el protocolo MIPv4

¹⁰ En el apartado 3.4.3 del Capítulo 3 se proporcionan más detalles de los protocolos IPsec y IKEv2.

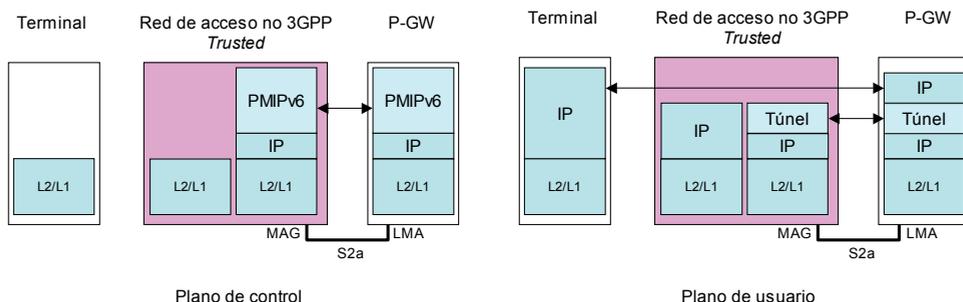


Figura 7.17 Arquitectura protocolos de la interfaz S2a en redes no 3GPP *trusted*

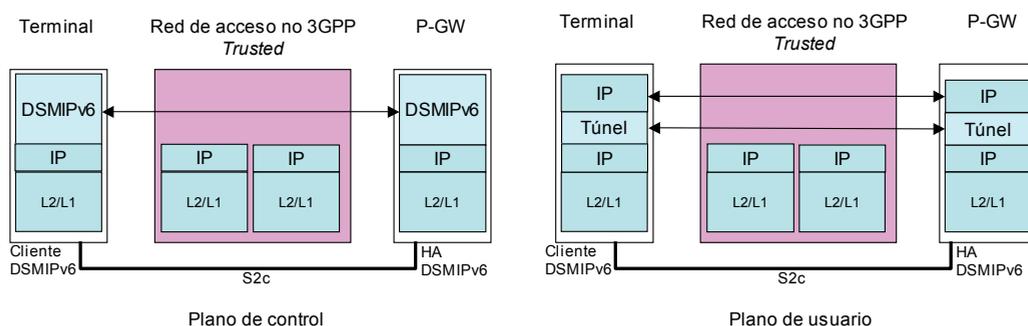


Figura 7.18 Arquitectura protocolos de la interfaz S2c en redes no 3GPP *trusted*

La solución de *interworking* para redes *trusted* también permite el control de las políticas de uso (e.g., control de QoS) y tarificación mediante la interconexión de la red de acceso no 3GPP con el subsistema PCC comentado en el apartado 3.2.4 del Capítulo 3. La interconexión con el subsistema PCC se realiza a través de la utilización de la interfaz Gx (utilizada también para los accesos a través de E-UTRAN y las otras redes 3GPP) y de una interfaz específica para el acceso a redes no 3GPP *untrusted* denominada Gxa. La interfaz Gxa funciona de forma análoga a la interfaz Gxx descrita en el apartado 3.2.4 del Capítulo 3 para el caso de que la pasarela S-GW y P-GW se conecten mediante PMIPv6. Así, la necesidad de disponer de la interfaz Gxa (además de Gx) radica en el hecho de que la interfaz entre la P-GW y la red de acceso no 3GPP tampoco soporta la implementación de servicios portadores. Por tanto, en caso de que la red de acceso no 3GPP permita el establecimiento de servicios portadores con parámetros de QoS, éstos terminan en la misma red de acceso y no se extienden hasta la pasarela P-GW. Por ello, la aplicación de la parte de las reglas PCC concerniente al control de QoS y *gating* se realiza a través de la interfaz Gxa. El resto de funciones relacionadas con la tarificación siguen realizándose en la P-GW, que recibe las reglas PCC a través de la interfaz Gx.

Finalmente, a modo ilustrativo, en la Figura 7.19 y Figura 7.20 se describen los procedimientos de registro a través de redes no 3GPP *trusted* utilizando, respectivamente, la interfaz S2a y S2c. A continuación se detallan los pasos más relevantes de ambos procedimientos.

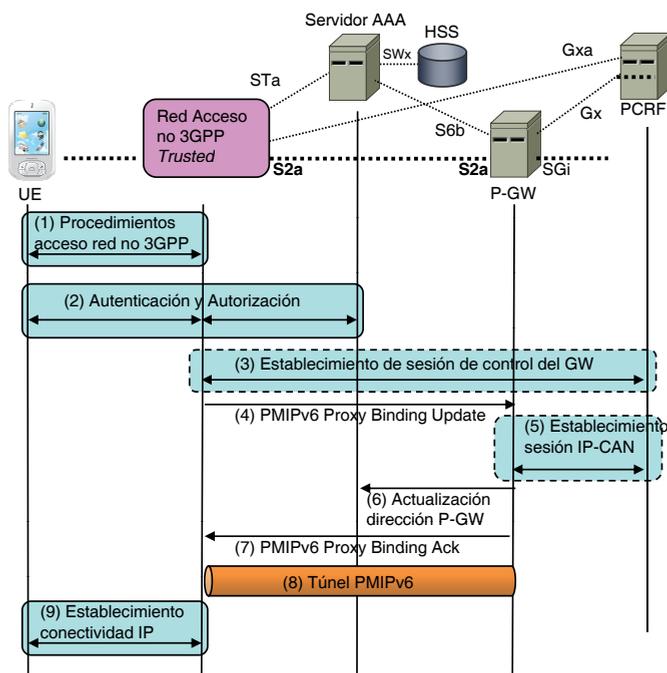


Figura 7.19 Procedimiento de registro en redes no 3GPP *trusted* utilizando la interfaz S2a

El procedimiento de registro mediante la interfaz S2a (Figura 7.19) consiste en:

(1) Acceso a la red no 3GPP

El terminal se conecta a la red de acceso no 3GPP mediante los mecanismos específicos de esta red.

(2) Autenticación y autorización

La autenticación y autorización del usuario se realiza mediante el protocolo EAP-AKA (*Extensible Authentication Protocol – Authentication and Key Agreement*) o su versión mejorada EAP-AKA'. El protocolo EAP-AKA es una adaptación del protocolo EAP¹¹ que permite reutilizar el mecanismo AKA propio de las redes 3GPP LTE y UMTS a través de redes de acceso no 3GPP. El protocolo EAP-AKA se especifica en la RFC 4187 [48] y su versión mejorada en RFC 5448 [49]. El protocolo permite que un terminal con una tarjeta USIM sea autenticado por el Servidor AAA de la red troncal EPC mediante la señalización EAP intercambiada a través de la red de acceso no 3GPP. Como resultado de la interacción entre la red de acceso no 3GPP y el servidor AAA, la red de acceso recibe información relativa al usuario (e.g., identificador único para el usuario, parámetros de QoS, etc.) así como información necesaria para la operación del protocolo de movilidad PMIPv6 (e.g., dirección de la pasarela P-GW que actuará como LMA

¹¹ En el apartado 7.2.5.1 se proporcionan más detalles del protocolo EAP

del usuario). A modo de ejemplo, tanto Mobile WiMAX como las redes Wi-Fi con certificación WPA (*Wireless Protected Access*) soportan mecanismos de autenticación basados en EAP.

(3) y (5) Establecimiento de la sesión IP-CAN

Tal como ocurre en el procedimiento de registro a la red LTE a través de E-UTRAN descrito en detalle en el apartado 3.2.5 del Capítulo 3, en caso de que se utilice el subsistema PCC para accesos no 3GPP, debe procederse al establecimiento de una sesión IP-CAN con la entidad PCRF correspondiente. Una sesión IP-CAN es una asociación entre la red que proporciona el servicio de conectividad IP y el sistema PCC en base a una dirección IP, una identidad de usuario y un identificador de red PDN (e.g., APN). En el contexto de una sesión IP-CAN, la entidad PCRF envía a la red que proporciona el servicio de conectividad el conjunto de reglas PCC aplicables al usuario (i.e., políticas de QoS y de tarificación). En el caso de redes no 3GPP, el establecimiento de la sesión IP-CAN requiere que previamente la red de acceso no 3GPP inicie el procedimiento de establecimiento de una sesión de control (paso (3) en la Figura 7.19) a través de la interfaz Gxc. A través de esta sesión de control la entidad PCRF puede enviar a la red de acceso no 3GPP la información relativa a las políticas de QoS para el control de los servicios portadores. Posteriormente, en el paso (5) de Figura 7.19, la pasarela P-GW iniciará la sesión IP-CAN que se vinculará en la entidad PCRF a la sesión de control establecida en el paso (3). De esta forma, la entidad PCRF conoce tanto la pasarela P-GW asociada a una sesión IP-CAN como la pasarela correspondiente de la red de acceso no 3GPP.

(4), (6), (7) y (8) *Binding update* y establecimiento del túnel PMIPv6.

El establecimiento de la asociación PMIPv6 entre el equipo de la red de acceso no 3GPP que alberga las funciones de MAG y la pasarela P-GW que actúa como LMA, empieza mediante el envío del mensaje "*PMIPv6 Proxy Binding Update*" que contiene, entre otros, un identificador del usuario (proporcionado a la red de acceso en el paso (2) por el servidor AAA) y el identificador APN de la red externa a la que debe proporcionarse el acceso. En este momento, la pasarela P-GW podría iniciar el establecimiento de la sesión IP-CAN comentada en el párrafo anterior. También, antes de responder a la red de acceso, la pasarela P-GW informa al servidor AAA de su identidad y la de la red externa a la que se conectará al usuario (paso (6)). Esto permite que la base de datos HSS contenga siempre información de cuál es la pasarela P-GW que está dando servicio a un usuario. Esta información, tal como se verá en el apartado 7.4.4, es fundamental para el soporte de *handover* entre sistemas.

La pasarela P-GW determina la dirección IPv4 y/o IPv6 que va a asignarse al terminal y responde a la solicitud de *Binding Update* de la red de acceso no 3GPP. En el intercambio de mensajes (4) y (7) se transporta la información necesaria (i.e., identificadores de túnel GRE) para que pueda quedar establecido un túnel GRE entre ambos extremos (paso (8)).

(9) Establecimiento de la conectividad IP

Finalmente, la red de acceso, mediante los mecanismos de asignación de direcciones de que disponga, procede a asignar la dirección IP enviada desde la pasarela P-GW al terminal.

Esta dirección es la dirección HoA que el terminal puede utilizar sin ningún tipo de encapsulado en la red de acceso en virtud del esquema de funcionamiento de PMIPv6.

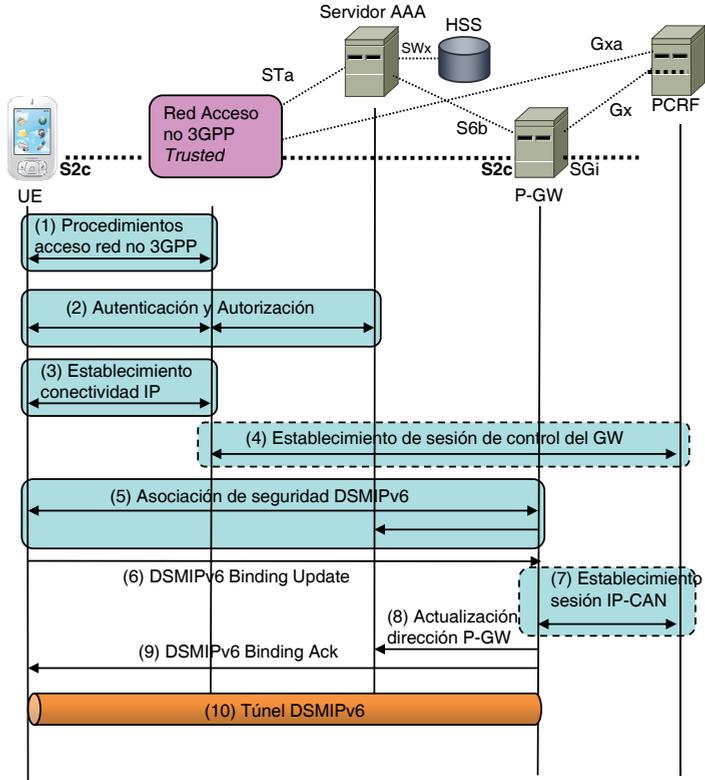


Figura 7.20 Procedimiento de registro en redes no 3GPP *trusted* utilizando la interfaz S2c

El procedimiento de registro mediante la interfaz S2c (Figura 7.19) es análogo al descrito para la interfaz S2a excepto en los siguientes aspectos:

(3) Establecimiento de la conectividad IP

La asignación de una dirección IP al terminal se produce una vez el usuario ha sido autenticado y autorizado. La dirección asignada permite que el terminal pueda utilizar el servicio de conectividad IP proporcionado por la red de acceso no 3GPP para comunicarse con la pasarela P-GW. Esta dirección constituye la dirección CoA utilizada por el protocolo DSMIPv6.

(5) Asociación de seguridad DSMIPv6

El uso de la interfaz S2c requiere que se establezca una asociación de seguridad entre el terminal y la pasarela P-GW sobre la que sustentan los intercambios de la señalización DSMIPv6. El establecimiento de la asociación de seguridad se sustenta en el protocolo IKEv2 [50] entre

terminal y P-GW y conlleva la interacción entre la pasarela P-GW y el servidor AAA en aras a autenticar al usuario. Nótese que esta segunda autenticación del usuario (la primera se realiza en el paso (2) de la Figura 7.19) no es necesaria en el procedimiento de registro mediante la interfaz S2c. Durante esta señalización, el terminal obtiene la dirección IP que utilizará como HoA.

(6), (8), (9) y (10) *Binding update* y establecimiento del túnel DSMIPv6

El terminal inicia la asociación DSMIPv6 mediante el envío del mensaje “*DSMIPv6 Binding Update*” donde se vincula la dirección HoA asignada durante el establecimiento de la asociación de seguridad DSMIPv6 en el paso (5) con la dirección CoA que dispone el terminal en la red de acceso no 3GPP. Tal como sucede en el procedimiento de registro para la interfaz S2a, la pasarela P-GW informa al Servidor AAA de los detalles del servicio de conectividad establecido para el usuario. El túnel DSMIPv6 para sustentar la transmisión del plano de usuario queda establecido una vez el terminal recibe la respuesta a su solicitud de *binding update*.

7.4.3 Solución genérica para redes Untrusted

La solución genérica de *interworking* en el caso de redes no 3GPP *untrusted* se muestra en la Figura 7.21. Los fundamentos de esta solución son los mismos que la solución presentada en el apartado anterior para redes *trusted*:

- Mediante la entidad de red denominada Servidor AAA (véase el apartado anterior para una descripción más detallada) se proporciona el soporte necesario a la red de acceso no 3GPP para llevar a cabo el control de acceso al servicio de conectividad IP de la red troncal EPC. En esta solución la interfaz equivalente a STa para redes *trusted* es ahora la interfaz SWa.
- La movilidad entre redes de acceso se fundamenta en el mantenimiento de la pasarela P-GW como punto de anclaje del servicio de conectividad, junto con la utilización de un protocolo de movilidad en la capa de red a través de la red no 3GPP. En este caso las interfaces asociadas con el soporte del protocolo de movilidad son S2b (basada en PMIPv6) y, al igual que en la solución para redes *trusted*, S2c (basada en DSMIPv6).
- La solución de *interworking* también permite el control de las políticas de uso y tarificación mediante la interconexión con el subsistema PCC a través de las interfaces Gxb y, al igual que en la anterior solución, Gx.

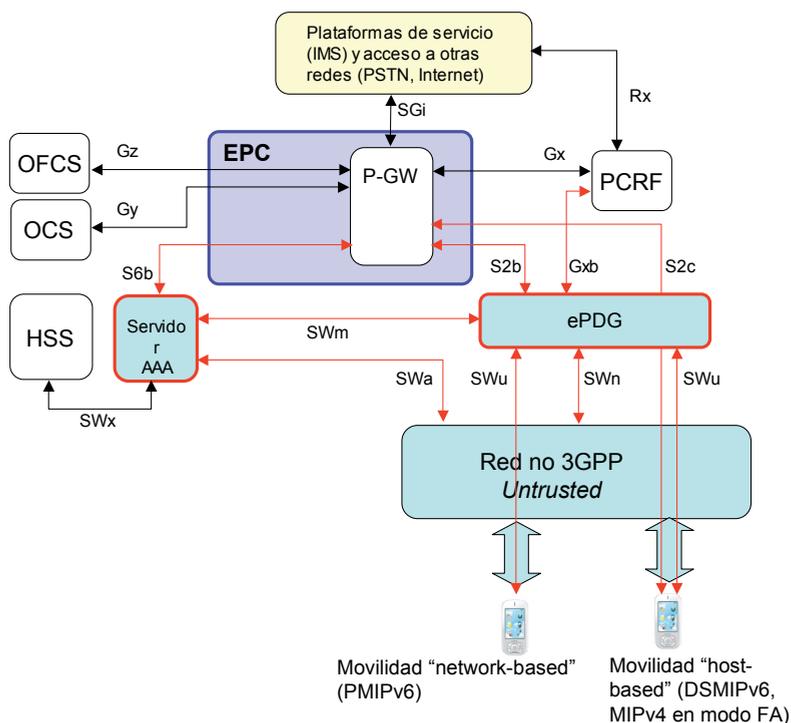


Figura 7.21 Solución genérica para redes no 3GPP *untrusted*

Por otro lado, la diferencia fundamental de esta solución con la especificada para redes no 3GPP *trusted* es la introducción de una nueva entidad de red denominada ePDG (*evolved Packet Data Gateway*). La finalidad principal del ePDG consiste en la provisión de un acceso seguro a los servicios de la red troncal EPC, dado que la red de acceso no 3GPP se considera que no proporciona las garantías de seguridad necesarias (de ahí la consideración como red *untrusted*). La provisión de dicho acceso seguro se materializa mediante la introducción de la interfaz SWu entre el terminal móvil y la pasarela ePDG. La interfaz SWu se basa en la utilización del protocolo IPsec y IKEv2 para el establecimiento de una asociación de seguridad entre terminal y ePDG de forma que toda la información transferida entre ambos elementos (señalización y datos de usuario) se transporta de forma segura a través de la red de acceso no 3GPP. La información de seguridad necesaria para el establecimiento de la asociación de seguridad (i.e., vectores de autenticación) se obtienen del Servidor AAA a través de la interfaz SWm, tal como se ilustra en la Figura 7.21. La solución genérica para redes *untrusted* es una extensión de la solución denominada *Interworking WLAN* (I-WLAN) introducida inicialmente en el contexto de redes UMTS [43] [44]. La solución I-WLAN permite que el usuario sea autenticado en una red WLAN utilizando las credenciales de una tarjeta SIM/USIM así como acceder a los servicios GPRS de la red UMTS (i.e., Nivel A y B de *interworking* según la clasificación del apartado 7.2.4). Sin embargo, la solución I-WLAN no soporta continuidad de servicio, limitación que ha sido solventada en la solución genérica para redes no 3GPP *untrusted*.

Las torres de protocolos de los planos de control y usuario entre el terminal y la pasarela P-GW se muestran en la Figura 7.22 y Figura 7.23 para, respectivamente, la opción de movilidad basada en el uso de la interfaz S2b y la basada en S2c. En el caso de la interfaz S2b el plano de control de la movilidad únicamente se extiende entre la pasarela ePDG (con funciones de MAG) y la pasarela P-GW (con funciones de LMA). Respecto al plano de usuario, la transmisión se realiza mediante la concatenación de dos túneles: un túnel IPsec entre el terminal y la pasarela ePDG y un túnel PMIPv6 entre la pasarela ePDG y la pasarela P-GW. En el caso de la interfaz S2c basada en DSMIPv6, el plano de control del protocolo de movilidad se extiende hasta el terminal y por ello es necesario proteger la señalización asociada mediante IPsec. En el plano de usuario, el establecimiento ahora del túnel de movilidad DSMIPv6 entre el terminal y la pasarela P-GW conlleva que, en la transferencia de información entre terminal y ePDG se produzca un doble encapsulado (el del túnel IPsec para proporcionar seguridad y el túnel de movilidad).

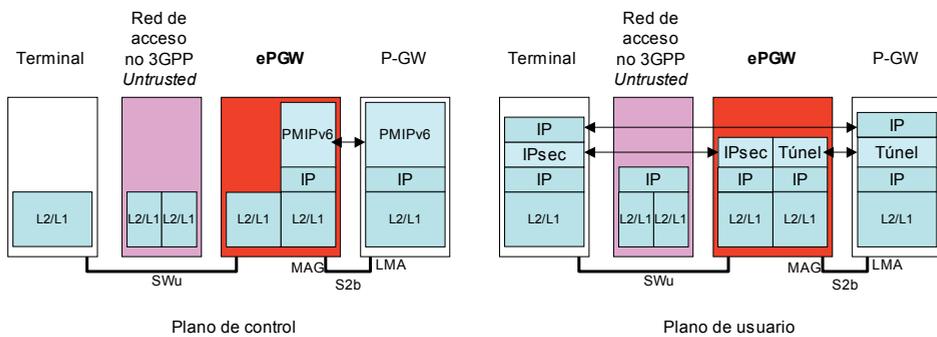


Figura 7.22 Arquitectura protocolos de la interfaz S2b en redes no 3GPP untrusted

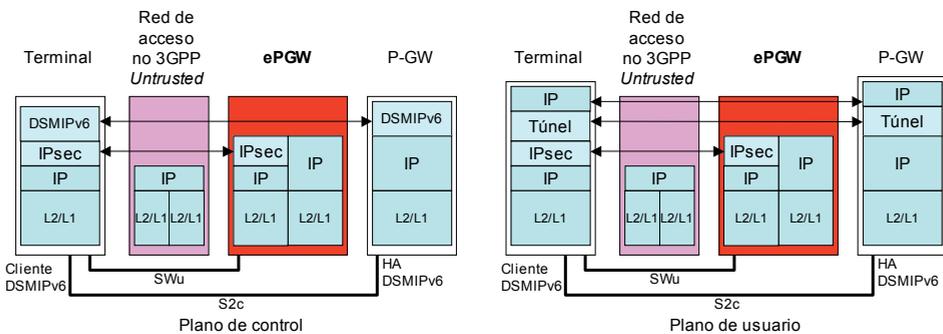


Figura 7.23 Arquitectura protocolos de la interfaz S2c en redes no 3GPP untrusted

Finalmente, análogamente a cómo se detalla en el apartado 7.4.2 para la solución con redes trusted, en la Figura 7.24 y Figura 7.25 se describen los procedimientos de registro a través de redes no 3GPP untrusted utilizando, respectivamente, la interfaz S2b y S2c. Dado que la mayoría de pasos en los que se estructuran los procedimientos para redes trusted y untrusted son similares, a continuación de detallan únicamente los aspectos en que los

procedimientos de registro en redes *untrusted* son diferentes a los de las redes *trusted* analizados en el apartado anterior.

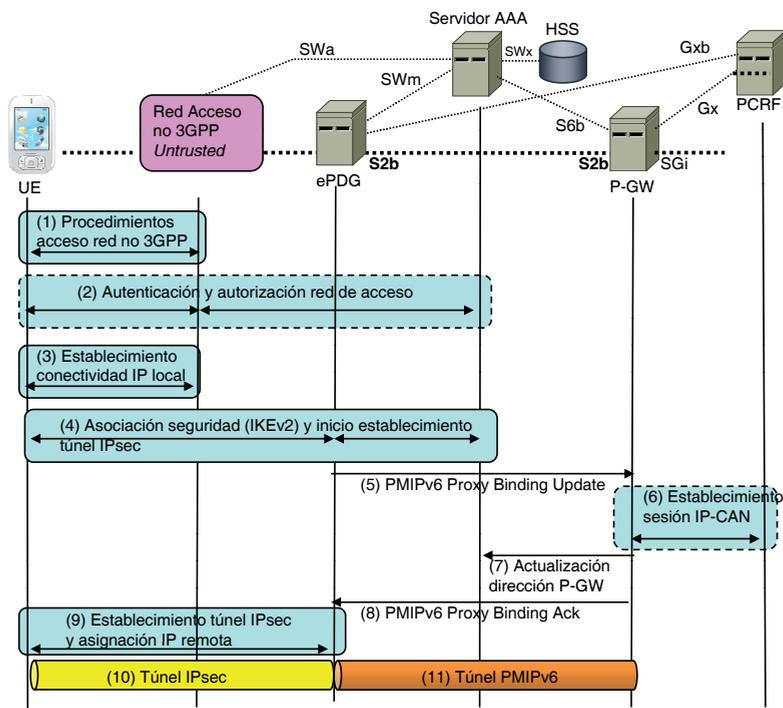


Figura 7.24 Procedimiento de registro en redes no 3GPP *untrusted* utilizando la interfaz S2b

Las principales diferencias del procedimiento de registro mediante la interfaz S2b en redes *untrusted* (Figura 7.24) y S2a en redes *trusted* (Figura 7.19) son:

- La autenticación del usuario en la red de acceso no 3GPP (paso (2)) es opcional y depende de la red de acceso en cuestión. El motivo es que en esta solución la autenticación y autorización de acceso al servicio de conectividad se realiza durante el establecimiento de la asociación de seguridad entre el terminal y la pasarela ePDG.
- El establecimiento de la conectividad IP local (paso (3)) se realiza de forma previa a la señalización PMIPv6 ya que es necesario disponer de una dirección IP válida con la que el usuario inicie el establecimiento de la asociación de seguridad IPsec con la pasarela ePDG (paso (4)). Esta dirección va a ser la utilizada como dirección externa (*outer address*) del túnel IPsec.
- El establecimiento de la asociación de seguridad se sustenta en el protocolo IKEv2 [50]. Durante este procedimiento (paso (4)), la pasarela ePDG recibe desde el servidor AAA la dirección de la pasarela P-GW contra la que iniciar la asociación de movilidad (i.e., PMIPv6 *binding update*).

- La asignación de la dirección IP HoA correspondiente al protocolo PMIPv6 se realiza en paralelo a la finalización del establecimiento del túnel IPsec (paso (9)). Así, en el túnel IPsec entre el terminal y la pasarela ePDG, la dirección interna del túnel (*inner address*) es la dirección IP HoA del terminal y la dirección externa es la dirección IP adquirida por el terminal en el paso (3).

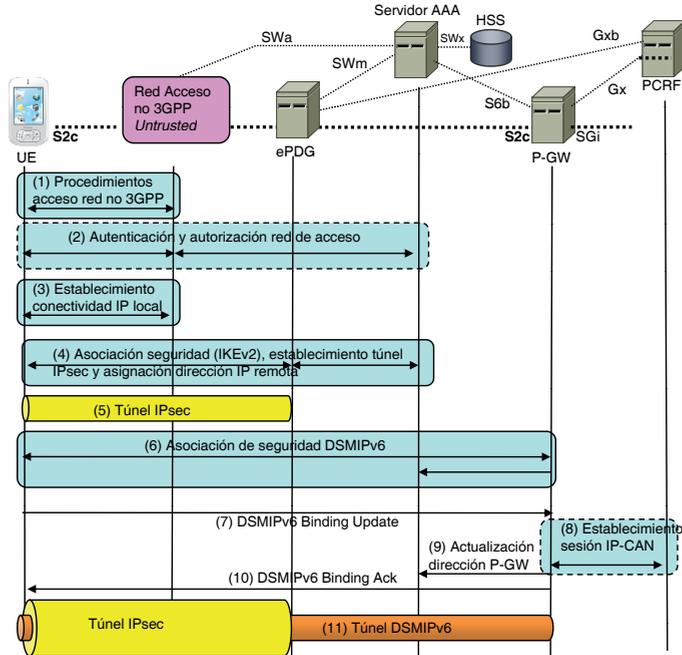


Figura 7.25 Procedimiento de registro en redes no 3GPP *untrusted* utilizando la interfaz S2c

En el caso de la interfaz S2c basada en DSMIPv6, las principales diferencias del procedimiento de registro entre redes *untrusted* (Figura 7.25) y *trusted* (Figura 7.20) son:

- La autenticación del usuario en la red de acceso no 3GPP (paso (2)) es opcional y depende de la red de acceso en cuestión, tal como sucede en el caso de la interfaz S2b en redes *untrusted*. En cambio, tras el establecimiento de la conectividad IP local en el paso (3), ahora es necesario que el terminal establezca una asociación de seguridad con la pasarela ePDG para poder acceder al servicio de conectividad.
- De forma previa a la señalización DSMIPv6, entre la pasarela ePDG y el terminal se establece un túnel IPsec. La dirección externa de este túnel es la dirección IP que el terminal adquiere en la red de acceso en el paso (3) y la dirección interna la proporciona la pasarela ePDG durante el paso (4).
- Establecido el túnel IPsec, el terminal inicia el establecimiento de la asociación de seguridad y del *binding update* del protocolo de movilidad DSMIPv6 de la misma forma

que en el caso de redes *trusted*. La única diferencia es que ahora toda la señalización se transporta sobre la red de acceso *untrusted* de forma protegida a través del túnel IPsec. Asimismo, la dirección que el terminal registra como CoA en la pasarela P-GW es la dirección IP interna del túnel IPsec (proporcionada en el paso (4)).

7.4.4 Movilidad entre sistemas

El sistema LTE soporta procedimientos de *handover* que permiten conseguir un Nivel C de *interworking* (véase el apartado 7.2.4) para las dos soluciones de *interworking* genéricas descritas en los apartados previos. Los procedimientos de *handover* soportados se basan en mantener la pasarela P-GW como punto de anclaje de las conexiones PDN cuando el terminal cambia de, por ejemplo, E-UTRAN a una red de acceso no 3GPP, así como cuando el cambio se produce entre dos redes de acceso no 3GPP. Los procedimientos pueden ser aplicados para cualquier tecnología de acceso y, por consiguiente, se trata de procedimientos genéricos que no están optimizados para ninguna red de acceso no 3GPP en concreto¹². Las características básicas de estos procedimientos pueden resumirse en tres aspectos fundamentales:

- La base de datos HSS siempre se encuentra actualizada con información que permite conocer cuál es la pasarela P-GW (o pasarelas) que proporcionan el servicio de conectividad IP a un usuario, independientemente de la red de acceso a través de la que se encuentre conectado¹³. Asimismo, en la base de datos se indica cuál es la entidad del plano de control a través de la cual se ha autorizado el acceso al servicio y que mantiene un contexto con los datos del usuario. Esta entidad es la entidad MME en el caso de E-UTRAN, SGSN para UTRAN y GERAN, y el Servidor AAA para las redes de acceso no 3GPP. Con toda esta información, la red troncal EPC puede gestionar y vincular entre sí las sesiones que el terminal pueda establecer a través de diferentes redes de acceso.
- La decisión de ejecutar un *handover* y el inicio del procedimiento corre a cargo del terminal. El principal motivo es que las soluciones de *interworking* genéricas no contemplan ninguna señalización específica que permita que, por ejemplo, el *handover* pueda ser iniciado desde la red (i.e., los mensajes de señalización de los protocolos NAS o RRC, utilizados por el terminal en el plano de control cuando está conectado a través de E-UTRAN, no contemplan ningún campo de información relacionado con, por ejemplo, redes IEEE 802). Asimismo, la red de acceso en la que se encuentre conectado el terminal no tiene por qué entender medidas de señales de otras tecnologías que le pudiera proporcionar el terminal de cara a realizar la decisión del traspaso en la red (i.e., el protocolo RRC de E-UTRAN no soporta el envío de medidas relativas al nivel de señal recibido en un punto de acceso o estación base IEEE 802, por tanto, un eNB no puede valorar ni ordenar el cambio de una conexión entre E-UTRAN y un sistema IEEE 802).

¹² En las especificaciones de 3GPP se hace referencia a dichos procedimientos de *handover* genéricos como procedimientos de *handover* no optimizados

¹³ En los procedimientos de registro ilustrados en las soluciones de *interworking* con redes no 3GPP (apartados 7.4.2 y 7.4.3) se ha visto como, para todas las variantes, la pasarela P-GW proporciona a la base de datos HSS, a través del servidor AAA, su identificador y el de la red externa a la que se conecta un terminal.

- No se contempla una fase de preparación del *handover* mientras el terminal se encuentre conectado a través de la red de acceso origen, a diferencia de lo que sucede en los mecanismos de *handover* entre redes 3GPP (véase el apartado 7.3.4.2). Por tanto, la consecución de tiempos de interrupción del servicio reducidos durante el cambio de red se sustenta básicamente en la posibilidad de hacer uso de terminales multimodo *dual-radio* que, a la vez que siguen cursando tráfico a través de la red de acceso origen, pueden llevar a cabo el establecimiento del servicio de conectividad a través de la red destino (e.g., autenticación, asociaciones de seguridad, *binding update* de los protocolos de movilidad, etc.).

En la Figura 7.26 se ilustran gráficamente los diferentes aspectos arriba expuestos. En el gráfico (a) de la Figura 7.26 se muestra una situación previa a la realización de un *handover* entre redes de acceso. En esta situación, el terminal se encuentra conectado a través de la red de acceso A (e.g., E-UTRAN) de forma que en la base de datos HSS se conoce la pasarela P-GW que proporciona el servicio y la entidad del plano de control asociada (e.g., MME en el caso de E-UTRAN). En esta situación el terminal detecta la red de acceso B (e.g., red Wi-Fi) y decide realizar el *handover* (cómo el terminal detecta la red de acceso B y qué criterios de decisión considera se analiza más adelante en este mismo apartado). Para ello, en el gráfico (b) de la Figura 7.26 se muestra como, en paralelo a la transferencia de información a través de la red de acceso A, el terminal inicia el procedimiento de registro y establecimiento del servicio de conectividad a través de la red B (e.g., autenticación en el punto de acceso Wi-Fi con EAP-AKA, obtención de conectividad local, etc.). La red troncal EPC detecta que el establecimiento del servicio de conectividad a través de la red B en realidad se trata de un *handover* de una sesión previamente establecida por el mismo usuario. Por este motivo, la red troncal EPC asigna a la nueva sesión la misma pasarela P-GW a través de la red de acceso B de forma que el terminal puede mantener la configuración IP (i.e., dirección IP) en la nueva red de acceso. Nótese que, si el terminal fuera *single-radio*, todo el procedimiento de registro y establecimiento de la conexión en la nueva red implicaría un tiempo de interrupción del servicio de conectividad IP. Por último, la pasarela P-GW realiza una conmutación de caminos para que el plano de usuario discurra por la nueva red de acceso, tal como se representa en el gráfico (c) de la Figura 7.26.

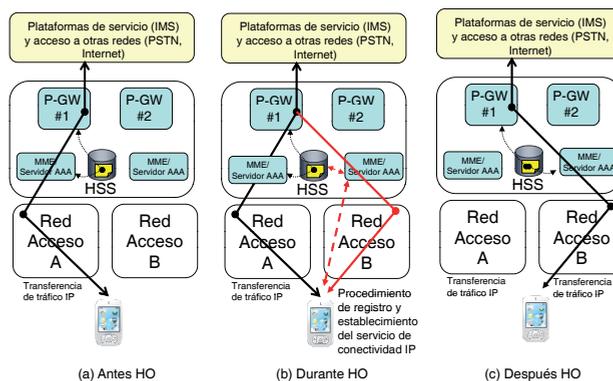


Figura 7.26 Esquema básico de movilidad entre redes

En la Figura 7.27 se describe de forma detallada un procedimiento de *handover* desde una red de acceso no 3GPP genérica hacia E-UTRAN. La situación de partida reflejada en la Figura 7.27 es la de un terminal que tiene establecido el servicio de conectividad IP a través de la red de acceso no 3GPP que utiliza la interfaz S2a o S2b para gestionar la movilidad. Así, el plano de usuario entre el terminal y la pasarela P-GW se sustenta en un túnel PMIPv6 establecido entre la pasarela P-GW y la red de acceso no 3GPP (en el caso de la interfaz S2b, a efectos de su representación en la Figura 7.27, se considera que el ePDG se encuentra englobado en el icono de "Red no 3GPP").

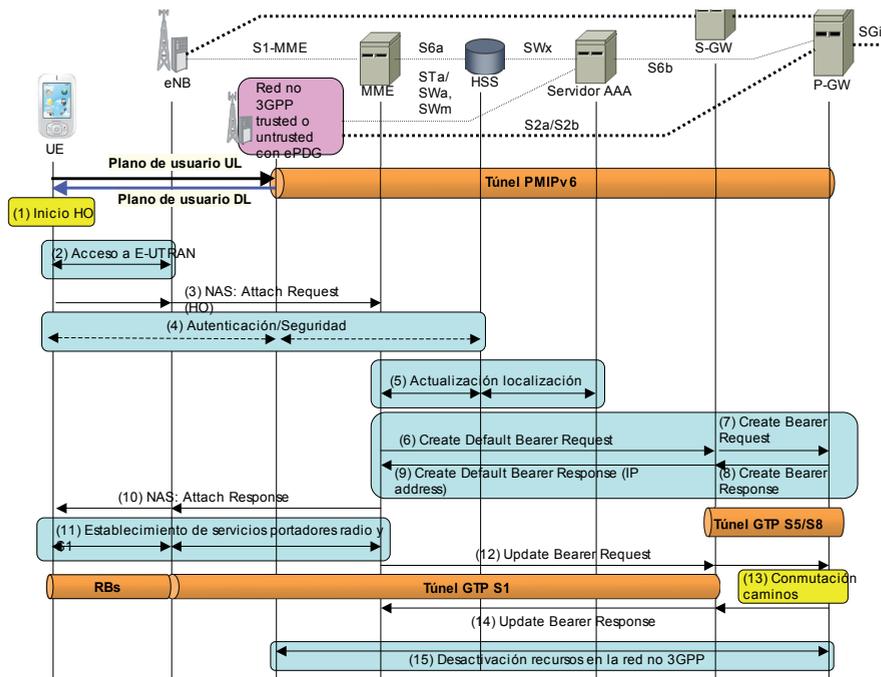


Figura 7.27 Procedimiento de handover entre una red de acceso no 3GPP genérica (*trusted* o *untrusted*) y E-UTRAN

Los diferentes pasos del procedimiento se describen a continuación.

(1) Inicio del handover

El terminal detecta la existencia de una celda E-UTRAN y decide continuar su servicio de conectividad IP a través de esa red. La detección de la nueva red puede ser consecuencia de que el terminal, o bien de forma periódica, o bien cuando la calidad de la red con la que mantiene la conexión lo requiera, realice búsquedas para detectar la presencia de señales piloto de otros posibles sistemas que operen en la vecindad. Para agilizar este proceso, el sistema 3GPP ha definido una entidad de red denominada como *Access Network Discovery and Selection Function (ANDSF)* que permite proporcionar a los terminales información

sobre qué redes alternativas tienen disponibles en su vecindad. El acceso a la entidad ANDSF desde los terminales se realiza mediante una señalización específica que se envía a través del propio servicio de conectividad IP. En la Figura 7.28 se muestra la interacción del terminal y la entidad ANDSF a través de la interfaz S14 definida a tal efecto y que se sustenta en el propio servicio de conectividad IP. Además de información con las características de las otras redes (e.g., frecuencias de operación, etc.), se contempla que la entidad ANDSF proporcione reglas de selección de red que los terminales consideren en sus decisiones de *handover*.

(2) - (5) Inicio del registro en E-UTRAN

Una vez tomada la decisión de realizar el *handover*, el terminal inicia el procedimiento de registro a través de E-UTRAN¹⁴. La diferencia principal con el procedimiento de registro inicial es que en el mensaje NAS de “*Attach Request*” (paso (3)) el terminal indica que se trata de un *handover*. De esta forma, la entidad MME sabe que el terminal tiene una sesión establecida a través de otra red y debe utilizar la información asociada (e.g., identificador de la pasarela P-GW y de la red externa APN) que exista en la base de datos HSS para llevar a cabo la selección de los nodos de la red troncal que van a sustentar el plano de usuario de la nueva conexión.

Si la entidad MME no dispone de un contexto establecido para el usuario, ésta procede a autenticarlo y actualizar su localización en la base de datos HSS. La actualización de la localización comporta también que la base de datos HSS cancele la localización previa del usuario en el servidor AAA utilizado para la provisión del servicio de conectividad en la red de acceso no 3GPP.

(6) – (9) Establecimiento del plano de usuario en la red troncal EPC

Con la información obtenida de la base de datos, la entidad MME inicia el establecimiento del servicio portador EPS por defecto de la interfaz S5/S8 mediante el envío del mensaje “*Create Default Bearer Request*” a la pasarela S-GW. El mensaje contiene la información necesaria para re-establecer el servicio portador EPS por defecto con la pasarela P-GW donde se encuentra anclada la conexión de la otra red. El mensaje enviado contiene una indicación de que se trata de un *handover*. Por este motivo, tras el establecimiento del túnel GTP en la interfaz S5/S8, la pasarela P-GW todavía no realiza la conmutación de caminos (i.e., el tráfico del usuario sigue fluyendo a través de la red no 3GPP). Asimismo, la dirección IP que tiene asignada el terminal en la conexión a través de la red no 3GPP se envía en toda esta señalización hacia la entidad MME.

Aunque no se ilustra en la Figura 7.27 por motivos de claridad, la creación del servicio portador EPS en la pasarela P-GW podría conllevar la modificación (o establecimiento) de una sesión IP-CAN con el subsistema PCC, de forma análoga a como se lleva a cabo durante el procedimiento de registro a través de las diferentes redes. A raíz de esta interacción la pasarela P-GW podría recibir reglas PCC que derivaran en el establecimiento de servicios portadores EPS dedicados.

¹⁴ El procedimiento de registro a través de E-UTRAN se detalla en el apartado 3.2.5.1 del Capítulo 3.

(10) y (11) Respuesta a la petición de registro y establecimiento de los servicios portadores en E-UTRAN

La entidad MME envía el mensaje NAS de “*Attach Response*” al terminal e inicia la activación de los servicios portadores (radio y S1) en el eNB donde se encuentra conectado el terminal. Más detalles de estos pasos pueden encontrarse en la descripción del procedimiento de registro a través de E-UTRAN.

(12) Finalización del establecimiento del servicio portador EPS y conmutación de caminos

Una vez establecidos los servicios portadores en el eNB, la entidad MME finaliza el establecimiento del túnel GTP en la interfaz S1 mediante el envío del mensaje indicado en el paso (12) que permite a la pasarela S-GW conocer la terminación del túnel en el eNB. En este punto el servicio portador EPS ya se encuentra establecido pero no operativo. Para ello, dado que la pasarela S-GW sabe que el procedimiento se ha derivado de una solicitud de *handover* (i.e., indicación de *handover* en el mensaje “*Attach Request*”), contacta con la pasarela P-GW para que proceda a realizar la conmutación del tráfico. A partir de este instante (paso (13)), todo el tráfico que reciba la pasarela P-GW va a encaminarse a través del servicio portador EPS establecido.

Nótese que, si el terminal fuera *single-radio*, la interrupción del servicio se extendería desde que el terminal activa su interfaz LTE (paso (2)) hasta la conmutación de caminos en la pasarela (paso (13)).

(15) Desactivación sesión en la red no 3GPP

Realizada la conmutación de caminos en la pasarela P-GW, ésta puede iniciar la señalización pertinente con la red de acceso no 3GPP. En el caso mostrado donde la interfaz es S2a o S2b, ambas basadas en PMIPv6, la pasarela P-GW envía el mensaje “*Binding Revocation*” que origina la desconexión del terminal en la red de acceso origen.

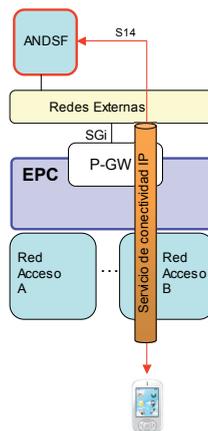


Figura 7.28 Funciones de soporte de la movilidad entre sistemas

La descripción completa de las diferentes combinaciones que existen de *handover* genérico con redes de acceso no 3GPP se proporciona en TS 23.402 [35].

7.4.5 Solución de interworking para redes CDMA2000

Además del marco genérico de movilidad, 3GPP ha especificado soluciones de *interworking* optimizadas para redes CDMA2000 [51] [52]. La necesidad de disponer de una solución optimizada para estas redes responde al hecho de que el sistema LTE se ha incorporado de forma clara en la hoja de ruta de muchos de los operadores con redes CDMA2000 [1]. Por este motivo, es importante que la introducción de redes de acceso LTE en las infraestructuras de redes CDMA pueda hacerse de forma gradual y garantizando la continuidad de servicio entre redes.

La solución optimizada para el servicio de transferencia por paquetes de una red CDMA2000 (denominado como *High Rate Packet Data*, HRPD) se basa en la solución genérica para redes *trusted* descrita en el apartado 7.4.2, complementada con dos interfaces nuevas denominadas S101 y S103 que se describirán a continuación. La solución de *interworking* resultante permite ofrecer un Nivel D de *interworking* para terminales *single-radio*. En la Figura 7.29 se muestra la arquitectura de la solución de *interworking*, remarcándose en trazo rojo fino las interfaces de la solución de *interworking* genérica y en trazo rojo grueso las nuevas interfaces de la solución optimizada para CDMA2000..

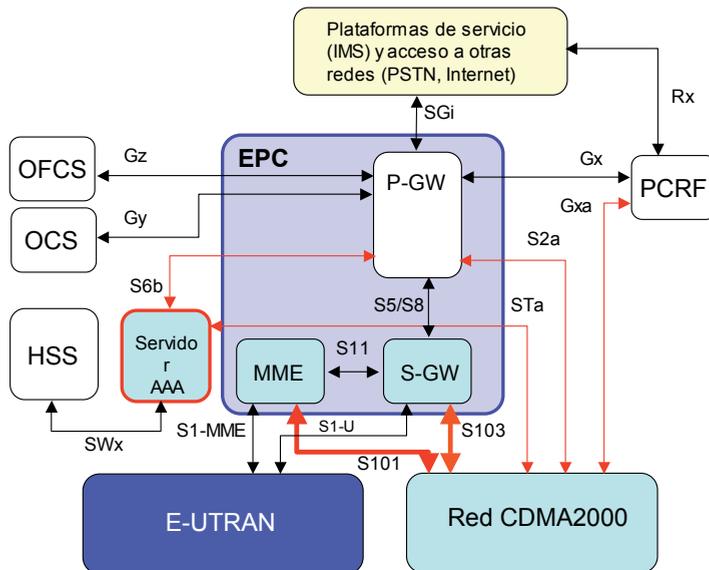


Figura 7.29 Solución de interworking optimizada para redes CDMA2000

La interfaz S101 es una interfaz de control concebida para poder transferir señalización de registro y de preparación del *handover* entre redes. Mediante esta interfaz, estando el ter-

minal conectado a la red CDMA2000, éste puede llevar a cabo la señalización de registro en E-UTRAN mediante la transferencia de la señalización NAS pertinente que se envía de forma encapsulada a través del enlace radio CDMA2000 y entre la red de acceso CDMA2000 y la entidad MME de la red troncal EPC a través de la interfaz S101. Este tipo de registro efectuado desde una conexión en otra red se denomina en las especificaciones 3GPP como pre-registro y puede realizarse en ambas direcciones. La finalidad última de este procedimiento es reducir al mínimo el tiempo de interrupción del servicio ocasionado por un procedimiento de traspaso. Nótese que en terminales *single-radio*, de no existir esta posibilidad, el procedimiento de registro debería hacerse necesariamente en la red destino de forma que, durante todo este proceso, el servicio estaría interrumpido (véase la descripción del *handover* genérico en el apartado anterior).

La interfaz S103 se utiliza para enviar tráfico de usuario entre la pasarela S-GW de la red troncal EPC y la red de acceso CDMA2000 durante la realización de un *handover*. En particular, el soporte de esta interfaz permite minimizar la pérdida de paquetes durante el proceso de *handover* desde E-UTRAN a la red CDMA2000 mediante el envío de los paquetes IP del usuario que pudieran quedar retenidos en la pasarela S-GW hacia la red de acceso destino.

Además de la introducción de las interfaces S101 y S103, la solución de *interworking* con redes CDMA2000 también ha supuesto la introducción de campos y procedimientos específicos en los protocolos RRC utilizados en el plano de control de la interfaz radio de los sistemas 3GPP. A modo de ejemplo, el protocolo RRC en E-UTRAN [45] incorpora mensajes específicos de información de sistema con datos de celdas CDMA2000 vecinas (e.g., mensaje SIB8 descrito en el anexo 2 del Capítulo 5), mensajes de configuración y envío de medidas relativas a esta tecnología, así como mensajes para ordenar la realización del procedimiento de *handover* a una celda CDMA2000 específica. Nótese que, a diferencia de la solución de *interworking* genérica, la incorporación de estas capacidades en el protocolo RRC permite que el procedimiento de *handover* se controle desde la red.

En la Figura 7.30 se muestra el procedimiento de *handover* en el caso de la solución de *interworking* optimizada para redes CDMA2000. La situación de partida es la de un terminal conectado a través de una red CDMA2000 HRPD de forma que el plano de usuario atraviesa la pasarela P-GW de la red troncal EPC (situación equivalente analizada en el apartado 7.4.4 para una red de acceso no 3GPP genérica que utiliza la interfaz S2a y PMIPv6).

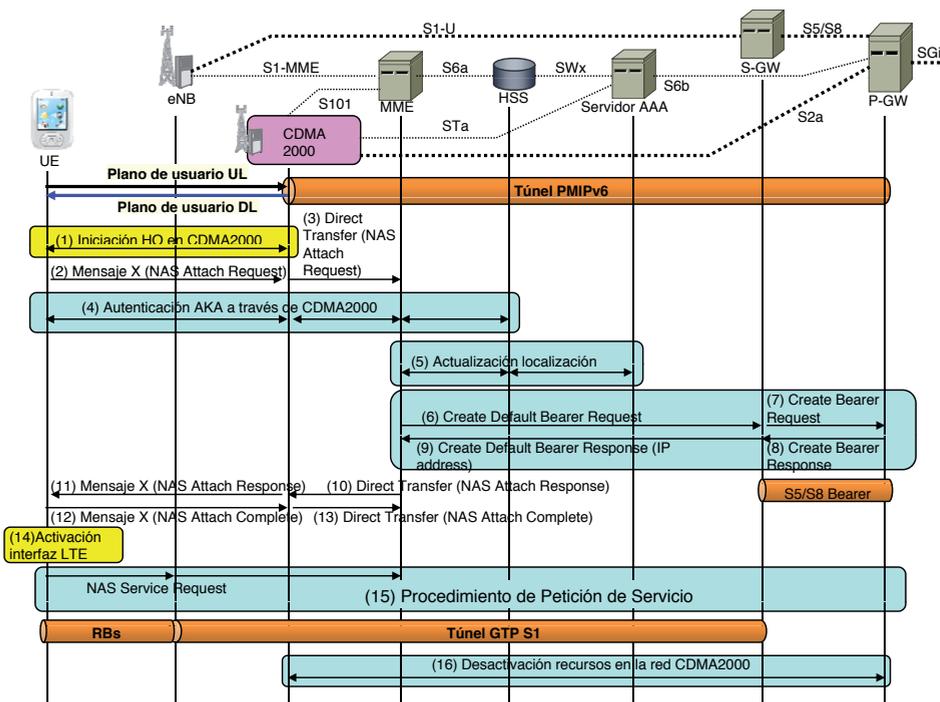


Figura 7.30 Solución de interworking optimizada para redes CDMA2000

Los pasos representados en la Figura 7.30 son los siguientes:

(1) Iniciación del *handover*

Los cambios mencionados anteriormente en los protocolos radio del plano de control de las redes implicadas permiten que el *handover* se controle desde la red. Así, en el ejemplo mostrado en la Figura 7.30, la decisión de cambiar a E-UTRAN podría realizarla la red de acceso CDMA2000 e indicar dicha decisión al terminal.

(2) - (5) Inicio del pre-registro

Tomada la decisión de cambiar a una celda E-UTRAN, el terminal inicia un procedimiento de pre-registro en la red troncal EPC. La característica principal de este procedimiento es que se realiza sin que el terminal deje de estar conectado a la red CDMA2000. Así, mediante la señalización radio propia del sistema CDMA2000, representada en la Figura 7.30 de forma genérica como "Mensaje X", el terminal envía la señalización NAS del sistema LTE con el mensaje "*Attach Request*". El mensaje NAS se transporta de forma encapsulada en la interfaz radio CDMA2000 y a través de la interfaz S101 entre ambas redes.

Tal como ocurre en el caso del procedimiento de *handover* genérico comentado en el apartado 7.4.4, la entidad MME puede obtener el identificador de la pasarela P-GW y la red externa APN a la que el usuario se encuentra actualmente conectado a partir de los datos

guardados en la base de datos HSS. Asimismo, la entidad MME autentica el usuario y procede a la actualización de su localización.

(6) – (9) Activación de los servicios portadores EPS en la red troncal EPC

El pre-registro comporta también la activación de los servicios portadores EPS en la red troncal EPC. Sin embargo, a diferencia de lo que ocurre en un registro convencional a través de E-UTRAN, el plano de usuario queda establecido únicamente entre las pasarelas P-GW y S-GW de forma equivalente a lo que sucede cuando un terminal se encuentra en estado *idle*.

(10) – (14) Activación de la interfaz LTE

El pre-registro termina con la señalización NAS indicada en los pasos (9)-(13). Mediante esta señalización el terminal ya dispone de un identificador temporal válido (i.e., S-TMSI) que le vincula a un contexto en una entidad MME de la red troncal EPC. El terminal también es informado de los servicios portadores EPS que se han establecido en la red troncal (activos entre la pasarela S-GW y P-W). Concluida esta señalización el terminal puede desactivar la interfaz CDMA2000 y activar la interfaz LTE.

(15) Procedimiento de petición de servicio

El re-establecimiento del servicio de conectividad en E-UTRAN se realiza de forma análoga al caso de que un terminal se encuentre en estado *idle* y pase a estado activo. El procedimiento de petición de servicio se detalla en el apartado 3.2.5 del Capítulo 3.

(16) Liberación de recursos en la red CDMA2000

Finalmente, al igual que ocurre en el procedimiento de *handover* genérico comentado en el apartado 7.4.4, la pasarela P-GW inicia la señalización pertinente con la red de acceso CDMA para desactivar los recursos utilizados por el terminal que ya se encuentra operativo en E-UTRAN.

Además de la solución optimizada para redes CDMA2000, el 3GPP ha llevado a cabo algunos estudios técnicos para el desarrollo de una solución optimizada para el caso de redes Mobile WiMAX y terminales *single-radio* [53]. En cualquier caso, las especificaciones actuales del sistema LTE (Release 9) no recogen todavía dicha posibilidad.

7.5 Referencias Bibliográficas

- [1] CDMA Development Group, "CDMA Migration Path to LTE", presentación disponible on-line en www.cdg.org.
- [2] 3GPP TR 25.913, "Requirements for evolved UTRA and evolved UTRAN".
- [3] IETF RFC 3748, "Extensible Authentication Protocol (EAP)".
- [4] IETF RFC 3588, "Diameter Base Protocol".
- [5] IEEE P802.11u/D8.0, Jul, 2009, Draft standard for Information Technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 7: *Interworking* with External Networks.
- [6] IETF RFC 2661, "Layer Two Tunneling Protocol "L2TP".
- [7] IETF RFC 4301, "Security Architecture for the Internet Protocol".
- [8] IETF RFC 4303, "IP Encapsulating Security Payload (ESP)".
- [9] IETF RFC 3344, "IP Mobility Support for IPv4".
- [10] IETF RFC 3024, "Reverse Tunneling for Mobile IP, revised".
- [11] IETF RFC 3775, "Mobility support in IPv6".
- [12] IETF RFC 5555, "Mobile IPv6 Support for Dual Stack Hosts and Routers".
- [13] IEEE Wireless Communications Magazine, número especial sobre "Architectures and protocols for mobility management in all-IP mobile networks", abril 2008.
- [14] Ki-Sik Kong et al., "Mobility management for all-IP mobile networks: mobile IPv6 vs. proxy mobile IPv6" IEEE Wireless Communications, vol.15, no.2, pp.36-45, April 2008.
- [15] IETF RFC 5213, "Proxy Mobile IPv6".
- [16] IETF Internet-Draft, draft-ietf-netlmm-pmip6-ipv4-support-17.txt, "IPv4 Support for Proxy Mobile IPv6", septiembre 2009.
- [17] Salsano, S. et al., "SIP-based mobility management in next generation networks" Wireless Communications, IEEE , vol.15, no.2, pp.92-99, April 2008.
- [18] IETF RFC 4068, "Fast Handovers for Mobile IPv6".
- [19] IETF RFC 4881, "Low-Latency Handoffs in Mobile IPv4".
- [20] IETF RFC 4988, "Mobile IPv4 Fast Handovers".
- [21] Dutta, A. et al., "Media-independent pre-authentication supporting secure interdomain *handover* optimization" IEEE Wireless Communications, vol.15, no.2, pp.55-64, April 2008.

- [22] IETF RFC C4067, "Context Transfer Protocol (CXTTP)".
- [23] IETF RFC 5113, "Network Discovery and Selection Problem".
- [24] IEEE Std 802.21-2008, "IEEE Standard for Local and Metropolitan Area Networks: Media Independent *Handover* Services."
- [25] Eastwood, L. et al., "Mobility using IEEE 802.21 in a heterogeneous IEEE 802.16/802.11-based, IMT-advanced (4g) network," IEEE Wireless Communications, vol.15, no.2, April 2008.
- [26] 3GPP TS 36.304 "User Equipment (UE) procedures in *idle* mode (Release 8)".
- [27] 3GPP TS 25.401 "UTRAN overall description".
- [28] 3GPP TS 43.051, "GSM/EDGE Radio Access Network; Overall description - Stage 2".
- [29] 3GPP TS 23.060, "General Packet Radio Service (GPRS); Service description; Stage 2".
- [30] 3GPP TS 23.401, "General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access".
- [31] 3GPP TS 22.173, "IP Multimedia Core Network Subsystem (IMS) Multimedia Telephony Service and supplementary services; Stage 1".
- [32] 3GPP TS 23.206, "Voice Call Continuity (VCC) between Circuit Switched (CS) and IP Multimedia Subsystem (IMS); Stage 2".
- [33] 3GPP TS 23.216, "Single Radio Voice Call Continuity (SRVCC); Stage 2 (Release 8)".
- [34] Magnus Olsson et al., "SAE and the Evolved Packet Core: Driving the Mobile Broadband Revolution", Academic Press, Elsevier, 2009, ISBN: 978-0-12-374826-3.
- [35] 3GPP TS 23.402, "Architecture enhancements for non-3GPP accesses".
- [36] WiMAX Network Forum Architecture, Stage 2, 3GPP-WiMAX *interworking*, Release 1, version 1.2, January 2008.
- [37] 3GPP TS 23.002, "Network Architecture".
- [38] 3GPP TS 29.273, "Evolved Packet System (EPS); 3GPP EPS AAA interfaces".
- [39] 3GPP TS 23.203, "Policy and charging control architecture".
- [40] 3GPP TS 29.275: "Proxy Mobile IPv6 (PMIPv6) based Mobility and Tunneling protocols; Stage 3".
- [41] 3GPP TS 24.303, "Mobility management based on Dual-Stack Mobile IPv6; Stage 3".
- [42] 3GPP TS 24.304, "Mobility management based on Mobile IPv4; User Equipment (UE) - foreign agent interface; Stage 3".
- [43] 3GPP TS 23.234, "3GPP system to Wireless Local Area Network (WLAN) *interworking*; System description".

- [44] Salkintzis, A.K.; Fors, C.; Pazhyannur, R., "WLAN-GPRS integration for next-generation mobile data networks", IEEE Wireless Communications, vol.9, no.5, October 2002.
- [45] 3GPP TS 36.331, "Evolved Universal Terrestrial Radio Access (E-UTRA), Radio Resource Control (RRC); Protocol specification".
- [46] 3GPP TS 29.060, "GPRS Tunnelling Protocol (GTP) across the Gn and Gp interface".
- [47] 3GPP TS 29.274 "3GPP Evolved Packet System (EPS); Evolved General Packet Radio Service (GPRS) Tunnelling Protocol for Control plane (GTPv2-C); Stage 3.
- [48] IETF RFC 4187, "Extensible Authentication Protocol Method for 3rd Generation Authentication and Key Agreement (EAP-AKA)".
- [49] IETF RFC 5448, "Improved Extensible Authentication Protocol Method for 3rd Generation Authentication and Key Agreement (EAP-AKA)".
- [50] IETF RFC 4306, "Internet Key Exchange Protocol Version 2".
- [51] 3GPP TS 29.276, "Optimized Handover Procedures and Protocols between E-UTRAN Access and cdma2000 HRPD Acces".
- [52] 3GPP TS 29.277, "Optimized Handover Procedures and Protocols between E-UTRAN Access and 1xRTT Access".
- [53] 3GPP TR 36.938 v8.0.0, "Improved Network Controlled Mobility between E-UTRAN and 3GPP2/Mobile WiMAX Radio Technologies", marzo 2008.

