



Universidad de Valladolid

Escuela Técnica Superior de
Ingenieros de Telecomunicación

Dpto. Teoría de la Señal y Comunicaciones e Ingeniería Telemática

TESIS DOCTORAL:

**CARACTERIZACIÓN DE LOS SERVICIOS OFRECIDOS EN
REDES UTRAN, MODELADO DE SU CONSUMO DE
RECURSOS Y ANÁLISIS DE SU GRADO DE SERVICIO.**

Presentada por Pablo Alonso García para optar al grado de
doctor por la Universidad de Valladolid

Dirigida por:
Alonso Alonso Alonso
Belén Carro Martínez
Javier Manuel Aguiar Pérez



Universidad de Valladolid

Escuela Técnica Superior de
Ingenieros de Telecomunicación

Dpto. Teoría de la Señal y Comunicaciones e Ingeniería Telemática

TESIS DOCTORAL:

**CARACTERIZACIÓN DE LOS SERVICIOS OFRECIDOS EN
REDES UTRAN, MODELADO DE SU CONSUMO DE
RECURSOS Y ANÁLISIS DE SU GRADO DE SERVICIO.**

Presentada por Pablo Alonso García para optar al grado de
doctor por la Universidad de Valladolid

Dirigida por:
Alonso Alonso Alonso
Belén Carro Martínez
Javier Manuel Aguiar Pérez

DEDICATORIA

A mis padres, por vuestro cariño, ejemplo y apoyo.

A Laura y al peque, por vuestro amor y paciencia.

A Maye y Pepín. Siempre os llevo en mi corazón.

A toda mi familia, con mucho afecto.

AGRADECIMIENTOS

Agradezco especialmente a Iván su ayuda durante toda la tesis.

Sin tu paciencia y buen hacer, no lo habría conseguido.

Alberto, tu ayuda también ha sido fundamental, y te la agradezco mucho.

Roberto, gracias por tu amistad, por preocuparte por mí y animarme siempre a hacerla.

Alonso, Belén y Javier, gracias por vuestra amistad, apoyo y consejo.

RESUMEN

En la actualidad, el dimensionamiento de la capacidad en los diferentes recursos de la red UTRAN se está realizando de forma empírica. Es decir, se amplía capacidad en los diferentes recursos y se observa la evolución de los fallos de conexión en el nodo. La presente Tesis propone un modelado de la gestión de la capacidad de los recursos de un nodo B ante la demanda simultánea de múltiples servicios. Este modelado permitirá dimensionar la capacidad de cada uno de los recursos para una combinación de tráficos multiservicio determinada, así como estimar los fallos de conexión achacables a carencias de capacidad en cada uno de los recursos. Ayudará asimismo a distinguir fallos de conexión debidos a congestión de fallos debidos a un mal funcionamiento hardware del recurso. El resultado de este trabajo será un modelo que permita a los operadores de red conocer la distribución de recursos necesaria para atender una determinada demanda de tráficos multiservicio. Así, se podrán corregir situaciones de exceso o defecto de recursos, consiguiendo un ahorro de coste operativo para el operador y mejorando la accesibilidad de la red. Se podrán detectar situaciones de accesibilidad degradada debido a fallos hardware y se podrá realizar el dimensionamiento óptimo de la red para una demanda futura estimada. Esta aplicación es de gran interés en la actualidad, donde es habitual que en grandes concentraciones de usuarios (conciertos, manifestaciones, encuentros deportivos, ...) la red no es capaz de atender la demanda generada. Ya no tanto en voz, donde el consumo de recursos es más liviano, sino en transferencia de datos y en señalización. Los canales de tráfico ascendente suelen estar más pobremente dimensionados y en ocasiones están provocando fallos para los smartphone que intentan subir ficheros pesados a un servidor (fotos, video). En el downlink hay un uso muy intensivo (video en streaming, por ejemplo), que se ve penalizado en casos de saturación. La señalización tiene un tráfico elevado al estar estos dispositivos continuamente pasando de estado *idle* a estado activo (por ejemplo cuando una aplicación tipo *twitter* se conecta recurrentemente para comprobar si hay mensajes nuevos).

Los servicios pueden clasificarse de la siguiente forma:

- ⇒ Garantizados (o rígidos): para admitir la conexión, la red comprueba si cuenta con los recursos suficientes, los reserva, y sólo los libera al finalizar la misma.
- ⇒ No garantizados (o elásticos): no se reservan recursos para una conexión, por lo que pueden ser desviados a otro servicio de mayor prioridad, reduciendo el Grado de Servicio. Adaptan su ocupación del recurso a la disponibilidad de dicho recurso de forma dinámica [2] mediante la participación de protocolos que implementan funciones tales como el control de flujo. Un mismo servicio puede consumir de forma elástica un recurso, mientras que consume de forma rígida otro. Así, por ejemplo, el ancho de banda de plano de usuario de Iub en las conexiones de datos se consume de forma elástica, mientras que el ancho de banda de plano de control empleado en la señalización se consume de forma rígida. Una correcta caracterización de los diferentes servicios permitirá modelar su consumo de recursos.

Una conexión consume fundamentalmente los siguientes recursos a nivel radio:

- ⇒ Ancho de banda determinado para el enlace entre el Nodo B y la RNC (enlace Iub). El ancho de banda disponible en cada momento dependerá de los diferentes servicios, ya que cada uno de ellos demanda una tasa binaria distinta.
- ⇒ Códigos de canalización, cuya característica más explícita son los códigos de expansión o Spreading Codes (SC). Este recurso es especialmente importante en el downlink, ya que

los códigos de canalización parten de un único árbol generador, mientras que en el enlace ascendente cada usuario tiene su propio árbol de códigos. Por ello, éste será un factor limitante únicamente en el enlace descendente.

⇒ Channel Elements (CE). Se define como una unidad de medida de los recursos de procesado en banda base por parte de los nodos B, implementados en el consumo de una conexión de voz.

En la presente Tesis, se utiliza el modelo Kaufman-Roberts para estimar la probabilidad de bloqueo para cada uno de los servicios que ha de atender un recurso. Se usa una versión modificada aprovechando la FFT, en lugar de la versión recursiva clásica, consiguiendo una notable mejora en el tiempo de computación. Ésto permite que el modelo pueda aplicarse como herramienta de trabajo habitual en el operador de red.

Las probabilidades de bloqueo por servicio se agrupan, obteniendo una probabilidad agregada en cada recurso. Posteriormente, se obtiene una probabilidad de bloqueo global, considerando la concatenación de recursos (y teniendo en cuenta los intentos de conexión que se bloquean al paso por cada uno de ellos). Por último, los resultados se validan en escenarios reales, contrastando los valores del modelo con los reportados por la red. La red reporta la probabilidad de bloqueo global del sistema, y los valores son similares a los obtenidos por el modelo. Pero este modelo permite obtener los valores a nivel de recurso, algo no disponible en los contadores reportados por la red.

El tráfico de entrada que usamos en el modelo es el reportado por la red, con la salvedad de que la red reporta cada 15 minutos. Y en la tesis se usan realiza un conformado de tráfico basado en una distribución de Poisson, que permite obtener valores con granularidad de 1 minuto, con un error muy bajo de reconstrucción.

El modelo presentado en esta Tesis está siendo usado satisfactoriamente en un operador de red móvil consiguiendo los siguientes beneficios:

- ⇒ Se consigue reducir el bloqueo, aumentando capacidad en el recurso limitante.
- ⇒ Se distribuye la capacidad de forma eficiente a los diferentes recursos, ampliando sólo los que son necesarios.
- ⇒ Se ha dimensionado correctamente la capacidad para eventos de alto tráfico, a partir de estimaciones de tráfico.
- ⇒ Se han detectado casos de fallo hardware, en los que éste, y no la congestión, era el motivo de los fallos de conexión.

ABSTRACT

Capacity management in mobile networks implies two main tasks: solving present blockage issues (detecting limiting resources, in order to upgrade them or differentiating congested from malfunctioning resources), and adequate resource provisioning to avoid blockage in a future event from which service demand is estimated (massively populated sports events, concerts, etc). Mobile network vendors do not provide operators with the tools to tackle these issues. UMTS network equipment reports Key Performance Indicators related to capacity shortage, such as failed connection attempts, but no clue is provided in order to detect the limiting resource or a subsystem malfunction, nor is any advice given on how many resources to provision in order to avoid blockage for a certain forecasted service demand. As an example, a high concentration of smartphones in a concert may increase uplink data traffic, requiring certain uplink resources to be improved. Modeling how each service demands capacity from different resources becomes crucial in order to adequately dimension the network.

In this Thesis I have modeled each resource behavior for a given service (voice, data, video call) demand. We obtain the overall accessibility and validate the model in real scenarios. We also propose how to detect when a specific resource is congested (or whether it is malfunctioning), and what resources must be upgraded to attend to a certain service traffic mix.

Current Universal Mobile Telecommunications System (UMTS) Terrestrial Radio Access Network (UTRAN), commonly referred to as 3G (3rd Generation Wireless Mobile Communication Technology), is the incumbent mobile technology in Europe. The authors have studied real 3G network scenarios in Spain, where, at the moment, Long Term Evolution / 4G (e-UTRAN) clusters are being deployed, but only in trial scenarios. So, not enough network KPIs are being collected in order to extend the dimensioning model to LTE. We have intentionally based our work on currently deployed 3G networks, where we have plenty of field experience to calibrate the model. I also propose how to extend this approach to future LTE based networks.

As stated by the 3GPP, services in UMTS are classified as Circuit Switched (CS) and Packet Switched (PS). CS are connection oriented services, such as voice and video telephony, while PS are data services, such as HSDPA, HSUPA and Release 99 (R99). CS Services are considered guaranteed traffic, while PS Services are considered non-guaranteed.

Uu is the radio interface between UE and Node B. Iub is the transmission interface from Node B to the RNC. CS and PS services demand resources at different Radio Access Network (RAN) levels: radio interface (spreading codes), baseband processing capacity (channel elements) and Iub capacity. PS and CS services share the consumption of channel elements and spreading code pool. Meanwhile, PS and CS services are carried by Iub_CS and Iub_PS respectively. PS traffic uses the remaining capacity in the Iub once the CS traffic has been allocated.

Whenever a resource shortage is detected, more resources need to be installed in the system. At radio level, capacity enhancements can be made by the addition of new carriers. Each 5MHz carrier contains a full Spreading Code (SC) tree. As stated by 3GPP, each service has a different Spreading Factor (SF) Code consumption. At Node B, processing capacity is measured in terms of Channel Elements (CE). One CE is the baseband processing capacity required in Node B to provide one voice channel, including the control plane signaling. Each service has a different demand of CEs. Increasing the processing capacity of Node B through additional CEs involves the installation of extra hardware baseband cards in Node B. At the Iub, capacity upgrades will imply increasing the number of Virtual Channel Connections (VCC) when using ATM, or increasing IP throughput

when using IP. Iub CS is usually implemented through legacy ATM Constant Bit Rate Virtual Circuits (CBR VC) while Iub PS is being migrated from ATM Variable Bit Rate VCs to IP.

Thus, three key points of constraint have been identified: radio, Node B and Iub capacity, hereinafter referred to as SC, CE and Iub respectively. A shortage of resources at any point implies a connection reject, degrading user's perceived quality. In consequence, being able to appropriately dimension UMTS resources at every subsystem is a must. Not in vain, insufficient resource allocation leads to service downgrade and over dimension is economically inefficient.

This work addresses the modeling of a UMTS scheme, considering resource availability at the three mentioned stages, i.e. SC, CE and Iub. These are the resources that can be upgraded by network operation. Other radio interface conditions, such as Uplink intercellular and traffic load interference level, or Downlink power consumption (depending on traffic and propagation loss) are not considered in this study. The impact of these factors can be diminished enhancing the number of Nodes in a specific area and reducing the Soft Handover Area (decreasing Nodes overlapping). But in real network operation, the former CE, SC and Iub are the main limiting resources, and its upgrading is part of the daily operation tasks. The difficulty for operators is to detect which one is the limiting resource and whether the accessibility is restricted due to congestion or to hardware failure. It is also crucial to know how to improve capacity in each resource (downlink/uplink CEs, SC, downlink/uplink Iub) in order to keep accessibility at acceptable levels for a forecast of services mix traffic (service traffic mix?).

The model described in this Thesis uses as input data the traffic for each service (CS and PS), and estimates the blocking probability for each resource. A system overall blocking probability (or its complementary, the accessibility) is then computed. The total combined accessibility is compared with the KPI reported accessibility in three different scenarios, now also including a known hardware impaired scenario to better demonstrate the capabilities of the model.

Services offered by the UMTS network under study are characterized in terms of their resource consumptions. Then, for each resource, a Kaufman-Roberts (K-R) based algorithm is applied using connection attempts, to produce the estimated accessibility for each service class (CS and PS). The accessibilities per resource will be combined to obtain an overall accessibility. Overall accessibility is available in network reported KPIs, though it is not available at a "per resource" level. The fact that the model does reach this detail is a key feature.

Índice general

1. Introducción	1
1.1. Introducción	1
1.2. Motivaciones	2
1.3. Antecedentes	2
1.3.1. Caracterización de los diferentes servicios	2
1.3.2. Consumo de recursos	3
1.3.3. Grado de servicio	4
1.4. Objetivos	5
1.5. Contribuciones de la Tesis	6
1.5.1. Caracterización de servicios y consumo de recursos	6
1.5.2. Degradación del grado de servicio	6
1.5.3. Validación de resultados en escenarios reales	7
1.6. Metodología	7
1.7. Estructura de la Tesis	8
2. Estado del arte	13
2.1. Estado del arte	13
3. Arquitectura de red UTRAN	17
3.1. Introducción	17
3.1.1. Primeros Pasos	17
3.1.2. Telefonía Celular	18
3.1.2.1. Primera Generación	19
3.1.2.2. Segunda Generación	19
3.1.2.3. Tercera Generación	20
3.1.3. Arquitectura de redes GSM	21
3.2. Redes UMTS. Arquitectura de red de acceso radio	22
3.3. Arquitectura de la red UTRAN	25
3.3.1. La controladora RNC	26
3.3.1.1. Papel lógico de la RNC	26
3.3.2. El Nodo B (estación base)	27
3.3.2.1. Estructura interna	27
3.3.2.2. Técnica de modulación	28
3.3.2.3. Funciones de control en el nodo B	28
3.3.3. Gestión de los recursos radio	29

3.4.	Arquitectura de la red UTRAN de Ericsson	29
3.4.1.	RNC	30
3.4.2.	Nodo B	30
3.5.	Tecnología LTE (<i>UMTS Long Term Evolution</i>)	32
3.5.1.	Introducción	32
3.5.2.	Arquitectura del sistema LTE	33
3.5.3.	Arquitectura de E-UTRAN	34
3.5.4.	Entidades de red: <i>Evolved Node B</i> (eNB)	36
3.6.	Comparativa E-UTRAN y UTRAN	37
3.7.	Descripción de las interfaces	39
3.7.1.	Interfaz Iub (Iu bis)	39
3.7.2.	Interfaz Iur	39
3.7.3.	Interfaz Iu	39
3.7.4.	Interfaz Uu	39
3.8.	Transmisión, capa física	40
3.8.1.	Radioenlaces	40
3.8.2.	xDSL	41
3.8.3.	Fibra Óptica	42
3.8.4.	Backhaul	42
3.9.	Transmisión, capa de red	42
3.9.1.	ATM	42
3.9.1.1.	Conceptos	43
3.9.1.2.	Calidad de Servicio	44
3.9.1.3.	Capas de Adaptación	45
3.9.2.	MPLS	45
3.9.2.1.	Beneficios de MPLS	49
3.10.	Interfaces de E-UTRAN	50
3.10.1.	Interfaz radio	51
3.10.2.	Interfaz eNB - EPC (S1)	52
3.10.3.	Interfaz eNB - eNB (X2)	54
3.11.	Comparativa E-UTRAN y UTRAN	54
3.12.	Conclusiones	56
4.	Modelado general KR sobre consumo de recursos	59
4.1.	Fórmula de Erlang-B	59
4.1.1.	Marco del modelo	59
4.1.2.	Limitación de la Fórmula de Erlang B	60
4.2.	Modelo de Kaufman-Roberts	60
4.2.1.	Marco del modelo	60
4.2.2.	Descripción del algoritmo multiservicio	61
4.2.3.	Ecuaciones	62
4.2.4.	Aspectos de la implementación	63
4.2.5.	Algoritmo de Kaufman revisado	64
4.2.6.	Mejora del algoritmo de Kaufman revisado basado en la FFT	64
4.3.	Comparación del rendimiento de los algoritmos	66

4.4. Conclusiones	66
5. Caracterización de servicios en una red UTRAN	69
5.1. Introducción	69
5.2. Servicios Portadores	69
5.3. Clases de Tráfico	71
5.3.1. Aplicaciones en tiempo real	71
5.3.1.1. Clase conversacional (Conversational class)	71
5.3.1.2. <i>Streaming Class</i> :	72
5.3.2. Aplicaciones NO tiempo real	72
5.3.2.1. <i>Interactive class</i> :	72
5.3.2.2. <i>Background class</i> :	72
5.4. Atributos de Calidad de Servicio	74
5.5. Mecanismos de Calidad de Servicio	74
5.5.1. Servicios sobre flujos garantizados	76
5.5.2. Servicios sobre flujos no garantizados	76
5.6. Conclusiones	76
6. Caracterización de recursos en una red UTRAN	79
6.1. Introducción	79
6.1.1. Requisitos de capacidad	81
6.2. Descripción de los recursos	83
6.3. Empleo de los recursos	83
6.3.1. Tráficos rígidos	84
6.3.2. Tráficos elásticos	84
6.3.2.1. Consideraciones sobre R99 NRT	85
6.3.2.2. Control de Admisión en R99 NRT	86
6.3.2.3. Cola WFQ	87
6.3.3. Tráfico HSPA	87
6.4. Códigos de canalización	87
6.5. Channel Elements	90
6.6. Iub sobre ATM	93
6.7. Iub sobre IP	94
6.7.1. Bloqueo en transmisión IP	94
6.7.2. Agregación y enrutado de Tráfico	95
6.7.3. Sentido del Bloqueo en IP	95
6.7.4. Dimensionamiento del Iub en IP	96
6.8. Estudio de las interferencias en el Uplink	98
6.8.1. Capacidad del Uplink y parámetro N_{pole}	98
6.8.2. Factor de carga en el Uplink	101
6.9. Estudio de la potencia en el Downlink	103
6.9.1. Factor de carga en el downlink	103
6.9.2. Estimación de la potencia total a transmitir	105
6.10. Conclusiones	105

7. Particularización modelo K-R al consumo de recursos	107
7.1. Introducción	107
7.2. Planteamiento del problema y aplicación multietapa	107
7.2.1. Marco general	108
7.2.2. Modelado del enlace descendente (DL)	111
7.2.2.1. Subsistema Iub (DL)	111
7.2.2.2. Subsistema CE (DL)	114
7.2.2.3. Subsistema SF (DL)	115
7.2.3. Combinación de factores en DL (IUB/CE/SF)	115
7.2.3.1. Modelo de bloques	115
7.2.3.2. Modelo probabilístico	116
7.3. Modelado del enlace ascendente	118
7.3.1. Subsistema CE (UL)	118
7.3.2. Subsistema Iub (UL)	120
7.3.3. Combinación de factores en UL (CE/Iub)	121
7.3.3.1. Modelo de bloques	121
7.3.3.2. Modelo probabilístico	121
7.4. Conformado de tráfico	122
7.4.1. Forma y parámetros de la distribución	123
7.4.1.1. Enfoque aproximativo desde el punto de vista infinitesimal	123
7.4.1.2. Particularización práctica desde el punto de vista discreto	125
7.4.1.3. Elección del vector de reparto (función de distribución discreta)	126
7.4.1.4. Reparto equitativo o distribución uniforme	126
7.4.1.5. Distribución normal	126
7.4.1.6. Distribución de Poisson	127
7.4.2. Implementación del conformado de tráfico en el modelo UTRAN	127
7.4.2.1. Cálculo del tráfico en Erlangs	127
7.4.2.2. Implementación del tráfico	129
7.5. Modelado de recursos	129
7.6. Cálculo de la capacidad del sistema	130
7.6.1. Capacidad del Iub	130
7.6.1.1. Nodos Dual-Stack	130
7.6.1.2. Nodos Full-IP	131
7.6.2. Capacidad para el recurso Códigos de Canalización (SF)	131
7.6.2.1. Detalles de implementación	132
7.6.3. Capacidad para el recurso Channel Elements (CE)	132
7.6.3.1. Enlace descendente	132
7.6.3.2. Enlace ascendente	134
7.7. Cálculo del bloqueo	134
7.8. Planteamiento de implementación del algoritmo de K-R para otros posibles factores	136
7.8.1. Implementación del algoritmo K-R para el análisis de la interferencia	136
7.8.2. Implementación del algoritmo K-R para el análisis de la potencia del amplificador	137
7.9. Procesado de estadísticas	137
7.10. KPIs	137

7.11. Elección Entorno de Implementación del Modelado	139
7.11.1. Opnet Modeler	139
7.11.2. Visual Basic for Applications	141
7.11.3. Programación sobre C# /Java/Python	142
7.11.3.1. C#	142
7.11.3.2. Java	142
7.11.3.3. Python	142
7.11.4. Mathworks MATLAB	143
7.12. Conclusiones	143
8. Comparativa entre los sistemas UTRAN modelado y real	147
8.1. Accesibilidad binaria y medida de calidad	147
8.1.1. Definición de calibración	147
8.1.2. Enfoque lógico o binario	147
8.1.3. Medida de calidad	148
8.2. Descripción de los casos de uso	150
8.2.1. Características generales de los casos de uso	152
8.2.2. Particularidades de los casos de uso	153
8.3. Calibración del tráfico	153
8.3.1. Elección del umbral de decisión	154
8.3.2. Modelado de tráfico como sistema realimentado o en lazo cerrado	155
8.3.2.1. Introducción a los sistemas realimentados	155
8.3.2.2. Particularización al modelo	156
8.3.2.3. Detalles de implementación	158
8.4. Evaluación de los resultados	161
8.4.1. Nodo «A»: nodo con capacidad IUB limitada (escenarios 1 y 2)	161
8.4.1.1. Escenario 1	161
8.4.1.2. Escenario 2	165
8.4.2. Nodo «B» (escenarios 3 y 4)	165
8.4.2.1. Escenario 3	169
8.4.2.2. Escenario 4	172
8.4.3. Nodo «C», ampliación del recurso CE (escenarios 5 y 6)	176
8.4.3.1. Escenario 5	176
8.4.3.2. Escenario 6	178
8.4.4. Nodo «D», configuración Full-IP (escenarios 7 y 8)	181
8.4.4.1. Escenario 7	181
8.4.4.2. Escenario 8	187
8.4.5. Nodo con fallo hardware	191
8.5. Conclusiones	193
9. Conclusiones y futuras líneas de trabajo	195
9.1. Conclusiones	195
9.1.1. Beneficios del uso de este modelado en la operación de redes UTRAN	196
9.1.2. Modelado multiservicio en los subsistemas UTRAN	196
9.1.3. Modelado y conformación del tráfico UTRAN	197

9.1.4.	Caracterización de los recursos	198
9.1.5.	Implementación del modelado	199
9.1.6.	Calibración del sistema vs. modelado del tráfico	199
9.2.	Líneas futuras	200
9.2.1.	Caracterización de servicios	201
9.2.2.	Consumo de recursos	201
9.2.3.	Detección del factor limitante	202
9.2.4.	Degradación en el GoS	203
9.2.5.	Aplicación de técnicas de aprendizaje	203
9.2.6.	Implementación de la herramienta sw de modelado	203
9.2.7.	Aplicación a informes de estadísticas automáticos	204
9.2.8.	Inclusión de otros factores limitantes de la red	204
A.	Código Fuente	205
A.1.	Análisis de la accesibilidad en un escenario determinado	205
A.1.1.	Rutina general	205
A.1.2.	Estudio de la accesibilidad	206
A.2.	Estudio de accesibilidades parciales	208
A.2.1.	Modelo accesibilidad Iub	208
A.2.1.1.	Nodos Dual-Stack	208
A.2.1.2.	Modelo accesibilidad nodos Full-IP	210
A.2.2.	Modelo accesibilidad channel elements	216
A.2.3.	Modelo accesibilidad códigos de canalización	217
A.2.4.	Combinación de accesibilidades parciales	219
A.2.5.	Digitalización de accesibilidad teórica	219
A.3.	Estudio de bloqueo en redes multiservicio	219
A.3.1.	Modelado de recursos CS	219
A.3.2.	Modelado de recursos PS	221
A.3.3.	Conformado de tráfico CS	221
A.3.4.	Conformado de tráfico PS	222
A.3.5.	Algoritmo original K-R	222
A.3.6.	Función Distribución K-R para el algoritmo original	223
A.3.7.	Función algoritmo K-R empleando FFT para evitar desbordamiento	224
A.3.8.	Función Erlang-B	225
A.3.9.	Función Erlanc-C	226
A.4.	Otras funciones	226
A.4.1.	Función constructor de árbol	226
A.4.2.	Actualización de tráficos entrantes a cada subsistema	228
A.4.2.1.	Tráficos CS	228
A.4.2.2.	Tráficos PS	228
A.4.3.	Reconstrucción de tráfico a intervalos de 15 minutos	228
A.4.4.	Carga de parámetros de los escenarios	229
A.4.5.	Calibrado de tráfico (parámetro λ)	231
A.4.6.	Representaciones gráficas	232

B. ATM (<i>Asynchronous Transfer Mode</i>)	235
B.1. Introducción	235
B.2. Conceptos	235
B.2.1. Celdas	235
B.2.2. Conexiones Virtuales	236
B.2.3. Conmutación en ATM	236
B.2.3.1. Conmutación de circuitos	236
B.2.3.2. Conmutación de paquetes	236
B.3. Calidad de Servicio	236
B.4. Capas de Adaptación	237
B.4.1. Capa AAL1	237
B.4.2. Capa AAL2	238
B.4.3. Capa AAL3-4	238
B.4.4. Capa AAL5	238
C. PDH-SDH	239
C.1. Introducción	239
C.2. PDH (<i>Plesiochronous Digital Hierarchy</i>)	239
C.3. SDH (<i>Synchronous Digital Hierarchy</i>)	240
C.3.1. Trama STM-1	240
D. Cálculo del ancho de banda medio usado en el RAB Iub ATM	243
D.0.2. Ancho de banda medio usado en el RAB Iub ATM	243
D.0.2.1. Señalización sobre ATM	243
D.0.2.2. Tráfico de voz sobre ATM	245
D.0.3. Algoritmo tráfico HSPA	249
E. Arquitectura de protocolos de UTRAN/E-UTRAN	251
E.1. Introducción	251
E.2. Arquitectura en capas	251
E.3. Protocolos de la interfaz Iub	253
E.4. Protocolos de la interfaz Iur	256
E.5. Protocolos de la interfaz Iu-CS	258
E.6. Protocolos de la interfaz Iu-PS	258
E.7. Protocolos radio	261
E.7.1. Protocolo RRC (Radio Resource Control)	262
E.7.1.1. Estados RRC	264
E.7.1.2. Procedimientos RRC	265
E.7.2. RLC	266
E.7.3. MAC	268
E.7.4. PDCP	268
E.8. Procedimientos UTRAN	269
E.8.1. «Paging»	269
E.8.2. Establecimiento de conexión RRC	270
E.8.3. Señalización móvil-red	272
E.8.4. Establecimiento de portadora de acceso radio (RAB)	273

E.8.5.	Trasposos suaves («Soft Handover»)	274
E.8.5.1.	Traspaso suave sobre un mismo RNC	274
E.8.5.2.	Traspaso suave entre RNC	275
E.8.6.	Reubicación de SRNS	275
E.9.	Protocolos E-UTRAN	276
E.9.1.	Protocolos en la interfaz radio	277
E.9.2.	Protocolos en las interfaces S1 y X2	278
E.9.3.	Plano de usuario entre UE y EPC	280
E.9.4.	Plano de control entre UE y EPC	280
F.	Notas sobre equipos Ericsson	283
F.1.	KPIs Reales	283
F.2.	Configuración de la red Ericsson	284
F.2.1.	Nodos con transmisión ATM	289
F.2.2.	Nodos con transmisión <i>DualStack</i>	289
F.2.3.	Nodos Full-IP	290
F.3.	Caracterización de recursos en equipos Ericsson	290
F.4.	Descripción de los casos de uso	291
F.4.1.	Características generales de los casos de uso	291
F.4.2.	Particularidades de los casos de uso	292
F.4.2.1.	Caso de uso A, CuadroF.3	292
F.4.2.2.	Caso de uso B CuadroF.4	292
F.4.2.3.	Caso de uso C CuadroF.5	292
F.4.2.4.	Caso de uso D CuadroF.6	292
G.	Técnicas de aprendizaje e inteligencia	293
G.1.	Introducción	293
G.2.	Modelado mediante árboles de decisión	294
G.2.1.	Concepto y estructura de un árbol de decisión	294
G.2.2.	Construcción de un árbol de decisión	295
G.2.3.	Entrenamiento del árbol	296
G.2.3.1.	Método de elección de la mejor variable discriminatoria	296
G.2.3.2.	Impureza de Gini	297
G.2.3.3.	Entropía	297
G.2.4.	Construcción recursiva del árbol de decisión	297
G.2.5.	Técnica de «poda»	298
G.2.6.	Trabajo con datos incompletos o perdidos	298
G.2.7.	Trabajo con salidas numéricas	299
G.2.8.	Cuándo utilizar árboles de decisión	299
G.2.9.	Detalles de implementación de árboles de decisión en la red UTRAN	300
G.3.	Redes Neuronales	302
G.3.1.	Concepto de Red Neuronal	302
G.3.2.	Aplicación de redes neuronales al modelado UTRAN	305
G.4.	Algoritmos genéticos	307
G.4.1.	Concepto de algoritmo genético	307

<i>ÍNDICE GENERAL</i>	XIX
G.4.2. Ventajas y desventajas. Limitaciones	307
G.4.3. Funcionamiento de los algoritmos genéticos	308
G.4.4. Aplicación de los algoritmos genéticos al modelado UTRAN	310
H. Material empleado	311
H.1. Material empleado	311
Glosario de acrónimos	313
Bibliografía	321

Índice de figuras

1.1. Ejemplo de tráficos garantizados y no garantizados	3
1.2. Proceso de validación	8
3.1. Evolución del nº de usuarios de telefonía móvil en los últimos 20 años	17
3.2. Concepto de red celular	18
3.3. Esquema de la arquitectura de una red GSM	20
3.4. Arquitectura general de UMTS	23
3.5. Elementos de la red en una PLMN	23
3.6. Arquitectura de red UTRAN	25
3.7. Arquitectura lógica básica de la RNC	26
3.8. Roles lógicos de las RNCs para una conexión UE-UTRAN	27
3.9. Estructura básica de un Nodo B	28
3.10. Técnicas de modulación utilizadas en WCDMA	29
3.11. Principales entidades en la gestión de recursos radio	30
3.12. Esquema básico de la arquitectura UTRAN	30
3.13. Esquema de componentes de un Nodo B Ericsson (RBS)	31
3.14. Arquitectura del sistema LTE	33
3.15. Red de acceso E-UTRAN	35
3.16. Comparativa de la arquitectura de red de E-UTRAN y UTRAN	37
3.17. Esquema de la red de transmisión del Iub	40
3.18. Funcionamiento de la modulación adaptativa	41
3.19. Celda ATM	43
3.20. Conexiones virtuales en ATM	43
3.21. Ejemplo de una red MPLS	48
3.22. Ilustración de los mecanismos de transferencia de información en la interfaz radio	52
3.23. Control de los servicios portadores radio y S1 a través de la interfaz S1-MME	53
3.24. Comparativa de la arquitectura de red de E-UTRAN y UTRAN	55
4.1. Marco del modelo de Kaufman-Roberts	60
4.2. Probabilidad de bloqueo servicio S_1	62
4.3. Conjunto Ω para el caso de servicios de voz, videollamada y señalización.	62
4.4. Comparación de la complejidad algorítmica de las diferentes implementaciones de Kaufman-Roberts (tiempos por iteración/muestra)	67
5.1. Arquitectura de servicios portadores UMTS	70
5.2. <i>Radio Access Bearer</i> (RAB) y señalización	71

5.3. Ganancia cualitativa según la diferenciación en QoS	74
6.1. Esquema de la gestión de tráfico	79
6.2. Definición de recursos en un sistema WCDMA	80
6.3. Esquema de la gestión de tráfico rígido	84
6.4. Esquema de la gestión de tráfico elástico	85
6.5. Esquema de la gestión de la solicitud de RAB R99 NRT	86
6.6. Esquema del funcionamiento del CAC de paquetes R99	87
6.7. Ejemplo de ensanchamiento de la señal en WCDMA	88
6.8. Proceso de ensanchamiento de la señal	89
6.9. Ensanchamiento de la señal	89
6.10. Ejemplo de árbol de códigos de ensanchamiento.	90
6.11. Árbol de códigos de cuatro capas	91
6.12. Consumo de recursos en tarjeta de nodo B	92
6.13. Cabeceras IPv4	94
6.14. Cabeceras IPv6	94
6.15. Agregado de tráfico	95
6.16. Enrutado de tráfico	95
6.17. Esquema del dimensionamiento de un salto de red en IP	96
6.18. Ejemplo de dimensionamiento IP	97
6.19. Conceptos de carga y capacidad en el Uplink	100
6.20. Incremento de la interferencia de celda con el nº de usuarios y sobrecarga	100
6.21. Incremento de la interferencia UL en función del factor de carga UL	102
6.22. Potencia transmitida por el nodo vs. factor de carga DL	105
7.1. Modelo (Primer nivel de desarrollo).	109
7.2. Comparación con KPIs reales	109
7.3. Modelo desarrollado (Nivel 2)	110
7.4. Modelo del enlace descendente (DL)	112
7.5. Modelo IUB en el downlink	112
7.6. Modelo CE en el downlink	114
7.7. Modelo SF en el downlink	115
7.8. Modelo de bloques para calcular la accesibilidad total CS	116
7.9. Combinación de factores en el DL.	117
7.10. Modelo para el enlace ascendente.	119
7.11. Modelo CE (UL)	119
7.12. Modelo Iub (UL)	120
7.13. Combinación de factores UL	121
7.14. Problema del eventanado de la función de densidad de probabilidad para una distribución no uniforme	124
7.15. Error en la reconstrucción de la muestra original para una distribución de Poisson eventanada	124
7.16. Ejemplo de interpolación de tráfico	125
7.17. Subsistema de conformado de tráfico	127
7.18. Esquema de cálculo de la accesibilidad por códigos de Spreading	133

7.19. Esquema para el cálculo del bloqueo debido a códigos de canalización	133
7.20. Subsistema cálculo de bloqueo para el algoritmo K-R	135
7.21. Arquitectura del procesado de estadísticas	138
7.22. Jerarquía de diseño Opnet	140
7.23. Proyecto sobre Opnet	141
7.24. Evolución de las métricas RMS y PPMCC para el caso de tráfico interpolado según distribución de Poisson.	144
7.25. Evolución de las métricas RMS y PPMCC para el caso de tráfico interpolado según distribución normal.	145
8.1. Umbral teórico de decisión de una accesibilidad teórica	149
8.2. Ejemplo de discretización y búsqueda de discrepancias	151
8.3. Esquema de comparación accesibilidades teórica y real	152
8.4. Escenario 1, $\Delta = f(\eta_t, \lambda)$	154
8.5. Tasa de discrepancias $\Delta = f(\eta_t, \text{escenario}), \lambda = 1$	155
8.6. Representación tridimensional de $\Delta = f(\eta_t, \lambda)$ en el esc. 1	155
8.7. Modelo general de un sistema realimentado	157
8.8. Particularización de modelo realimentado para la calibración del tráfico	158
8.9. Marco de implementación	159
8.10. Procedimiento para la verificación de calidad vs. λ	160
8.11. Tráfico CS durante un día de observación (nodo A, antes)	162
8.12. Tráfico PS durante un día de observación (nodo A, antes)	163
8.13. Accesibilidades real y teórica para tráfico CS en el nodo A.	163
8.14. Accesibilidades real y teórica para tráfico PS en el nodo A.	164
8.15. Predicción de la degradación en modelos teórico y real (nodo A, antes)	164
8.16. Ampliación de recursos en el nodo «A»	165
8.17. Tráfico CS durante un día de observación (nodo A, después)	166
8.18. Tráfico PS durante un día de observación (nodo A, después)	166
8.19. Accesibilidades real y teórica para el tráfico CS en el nodo A después de incrementar el recurso IUB.	167
8.20. Accesibilidades real y teórica para el tráfico PS en el nodo A después de incrementar el recurso IUB	168
8.21. Predicción de la degradación en modelos teórico y real (nodo A, después)	168
8.22. Tráfico CS durante un día de observación (nodo B, antes)	169
8.23. Tráfico PS durante un día de observación (nodo B, después)	170
8.24. Accesibilidades real y teórica para el tráfico CS en el nodo B (antes)	170
8.25. Accesibilidades real y teórica para el tráfico PS en el nodo B (antes)	171
8.26. Predicción de la degradación en modelos teórico y real (nodo B, antes)	172
8.27. Ampliación de recursos en el nodo «B»	173
8.28. Tráfico CS durante un día de observación (nodo B, después)	173
8.29. Tráfico PS durante un día de observación (nodo B, después)	174
8.30. Accesibilidades real y teórica para el tráfico CS en el nodo B (después)	174
8.31. Accesibilidades real y teórica para el tráfico PS en el nodo B (después)	175
8.32. Predicción de la degradación en modelos teórico y real (nodo B, después)	176
8.33. Tráfico CS durante un día de observación (nodo C, antes)	177

8.34. Tráfico PS durante un día de observación (nodo C, después)	177
8.35. Accesibilidades real y teórica para el tráfico CS en el nodo C (antes)	178
8.36. Accesibilidades real y teórica para el tráfico PS en el nodo C (antes)	179
8.37. Predicción de la degradación en modelos teórico y real (nodo C, antes)	179
8.38. Ampliación de recursos en el nodo «C»	180
8.39. Tráfico CS durante un día de observación (nodo C, después)	180
8.40. Tráfico PS durante un día de observación (nodo C, después)	181
8.41. Accesibilidades real y teórica para el tráfico CS en el nodo C (después)	182
8.42. Accesibilidades real y teórica para el tráfico PS en el nodo C (después)	183
8.43. Predicción de la degradación en modelos teórico y real (nodo C, después)	183
8.44. Tráfico CS durante un día de observación (nodo D, antes)	184
8.45. Tráfico PS durante un día de observación (nodo D, antes)	184
8.46. Accesibilidades real y teórica para el tráfico CS en el nodo D (antes)	185
8.47. Accesibilidades real y teórica para el tráfico PS en el nodo D (antes)	186
8.48. Predicción de la degradación en modelos teórico y real (nodo D, antes)	187
8.49. Ampliación de recursos en el nodo «D»	188
8.50. Tráfico CS durante un día de observación (nodo D, después)	188
8.51. Tráfico PS durante un día de observación (nodo D, después)	189
8.52. Accesibilidades real y teórica para el tráfico CS en el nodo D (después)	189
8.53. Accesibilidades real y teórica para el tráfico PS en el nodo D (después)	190
8.54. Predicción de la degradación en modelos teórico y real (nodo D, antes)	190
8.55. Tráfico CS durante un día de observación (nodo con fallos HW)	191
8.56. Tráfico PS durante un día de observación (nodo con fallos HW)	192
8.57. Accesibilidades real y teórica para tráfico CS en nodo con fallo hardware	192
8.58. Accesibilidades real y teórica para tráfico PS en nodo con fallo hardware	193
9.1. Compromiso en la elección del parámetro λ para una distribución de Poisson	200
9.2. Aplicación multietapa del algoritmo de K-R en el <i>downlink</i> ampliado	201
9.3. Aplicación multietapa del algoritmo de K-R en el <i>uplink</i> ampliado	202
B.1. Celda ATM	235
B.2. Conexiones virtuales en ATM	236
C.1. Estructura de trama STM-1	240
C.2. Mapeado de E1s sobre STM-1	241
D.1. Estructura de la trama DCH	243
D.2. SRB ATM UL Ancho de banda medio.	244
D.3. SRB ATM DL Ancho de banda medio.	245
D.4. SID UL ATM Ancho de banda medio, en tramas mudas.	246
D.5. SID UL ATM Ancho de banda medio, en tramas con muestra de voz.	247
D.6. SID DL ATM Ancho de banda medio, en tramas mudas.	248
D.7. SID DL ATM Ancho de banda medio, en tramas con muestra de voz.	248
D.8. Tráfico sobrante destinado a HSPA	249
E.1. Arquitectura de protocolos de UTRAN	252

E.2. Arquitectura general de protocolos de las interfaces UTRAN	252
E.3. Arquitectura de protocolos de la interfaz Iub	254
E.4. Transporte de tramas radio sobre la interfaz Iub	254
E.5. Funcionamiento del protocolo AAL2	255
E.6. Establecimiento y liberación de miniconexiones AAL2	256
E.7. Impacto de los trasposos suaves (Soft-HO) en UTRAN	256
E.8. Arquitectura de protocolos de la interfaz Iur	257
E.9. Arquitectura de protocolos de la interfaz Iu-CS	258
E.10. Arquitectura de protocolos de la interfaz Iu-PS	259
E.11. Encapsulado de paquetes de usuario en la interfaz Iu-PS	260
E.12. Comparación de los planos de usuario de GPRS y de UMTS	260
E.13. Transporte de datos y señalización móvil-red sobre UTRAN	261
E.14. Arquitectura de protocolos de la interfaz radio	262
E.15. SRB, RAB y conexión de señalización	263
E.16. Procedimientos de aviso («Paging»)	269
E.17. Procedimiento de establecimiento de conexión RRC	270
E.18. Intercambio de señalización móvil-red sobre la UTRAN	272
E.19. Procedimiento de establecimiento de RAB	273
E.20. Establecimiento de portadoras de datos AAL2	274
E.21. Traspaso suave entre Nodos B controlados por el mismo RNC	274
E.22. Traspaso suave entre Nodos B controlados por distintos RNC	275
E.23. Reubicación de SRNS	276
E.24. Protocolos de la interfaz radio de E-UTRAN	278
E.25. Protocolos en las interfaces S1 y X2	279
E.26. Protocolos del plano de usuario en E-UTRAN	280
E.27. Protocolos del plano de control en E-UTRAN	281
F.1. Ejemplo del KPI 3G-ERI-RAB-Att-CS.	284
F.2. Ejemplo del KPI 3G-ERI-RAB-Att-PS-R99	285
F.3. Ejemplo del KPI 3G-ERI-RAB-Att-HSDPA.	285
F.4. Ejemplo del KPI 3G-ERI-RAB-Att-HSUPA.	285
F.5. Ejemplo del KPI 3G-ERI-RAB-Fail-CS.	285
F.6. Ejemplo del KPI 3G-ERI-RAB-Fail-PS-R99	286
F.7. Ejemplo del KPI 3G-ERI-RAB-Fail-HSDPA.	286
F.8. Ejemplo del KPI 3G-ERI-RAB-Fail-HSUPA.	286
F.9. Ejemplo del KPI 3G-ERI-Num-USER-hsdpa	286
F.10. Ejemplo del KPI 3G-ERI-Num-USER-hsupa.	287
F.11. Ejemplo del KPI 3G-ERI-Minutes-Voice-q.	287
F.12. Ejemplo del KPI 3G-ERI-KB-PS-R99.	287
F.13. Ejemplo del KPI 3G-ERI-KB-HSDPA-DL-P6	288
F.14. Ejemplo del KPI 3G-ERI-KB-HSUPA.	288
F.15. Ejemplo del KPI 3G-ERI-SHO-factor	288
F.16. Ejemplo del KPI 3G-ERI-CSSR-CS	288
F.17. Ejemplo del KPI 3G-ERI-CSSR-PS	289

G.1. Ejemplo de árbol de decisión	295
G.2. Buffer o ventana de observación de accesibilidades reportadas	301
G.3. Estructura básica de una neurona artificial	303
G.4. Alternativas a la función $f(s)$ para una neurona artificial básica	304
G.5. Representación multicapa de una red neuronal	304
G.6. Ejemplo de red neuronal multicapa con n entradas y n salidas	304
G.7. Esquema básico de entrenamiento de una red neuronal	305
G.8. Modelado de red UTRAN utilizando redes neuronales	306
G.9. Esquema de entrenamiento de red neuronal UTRAN	306
G.10. Definición general de los individuos de un algoritmo genético	309
G.11. Operaciones de cruce de poblaciones y mutación de descendientes	309
G.12. Aplicación de algoritmos genéticos para modelado UTRAN	310

Índice de cuadros

3.1. Anchos de banda medio en el canal Iub	31
3.2. Capacidad radioenlaces Nokia-Siemens FlexiPacket vs. ancho de banda del canal y modulación empleados	41
3.3. Capacidad radioenlaces Huawei PTN vs. ancho de banda del canal y modulación empleados	42
3.4. Interfaces de E-UTRAN	50
5.1. Resumen tipos de tráfico en UMTS según QoS	73
5.2. Parámetros de QoS para tráficos en UMTS	73
5.3. Atributos de QoS del servicio portador RAB	75
5.4. Funciones de gestión de QoS en UMTS	75
6.1. Ejemplo de supuestos de tráfico utilizados durante una planificación inicial	81
6.2. Consumo de CE en tarjeta DL	92
6.3. Consumo de CE en tarjeta UL	92
6.4. Multiplexación PDH en Europa	93
6.5. Definición de parámetros para el cálculo del Factor de Carga Uplink η_{UL}	102
6.6. Valores objetivos típicos para la relación E_b/N_0 en el dimensionamiento de redes UMTS	103
6.7. Definición de parámetros para el cálculo del Factor de Carga Uplink η_{UL}	104
7.1. Tráficos entrante y saliente en modelo downlink	116
7.2. Tráficos en modelo UL	121
7.3. Tráficos CS de entrada utilizados en cada subsistema	128
7.4. Cálculo de tráfico en Erlangs	128
7.5. Modelado de recursos	130
8.1. Situaciones posibles durante la calibración del sistema	148
8.2. Definición de la variable lógica A_r	148
8.3. Definición de la variable lógica A_r	148
8.4. Operación XOR para la detección de discrepancias	149
8.5. Recursos para cada escenario	153
C.1. Tasas binarias para los niveles PDH europeos	240
C.2. Tasas binarias SDH	240
D.1. Anchos de banda medio en el canal Iub	249

E.1. Principales funciones RRC	263
E.2. Servicios que proporciona RRC	264
F.1. Descriptores de tráficos de los servicios sobre ATM	289
F.2. Configuración DualStack.	290
F.3. Recursos para cada escenario	292
F.4. Recursos para cada escenario	292
F.5. Recursos para cada escenario	292
F.6. Recursos para cada escenario	292
G.1. Ejemplo de variable discriminatoria ineficaz	296
G.2. Ejemplo de variable discriminatoria eficaz	296
G.3. Formato y ejemplo de casos de entrenamiento para el árbol de decisión	302
G.4. Atributos de salida para el árbol de decisión	302

Capítulo 1

Introducción

1.1. Introducción

Caracterización de los diferentes servicios

Las redes de comunicaciones móviles UTRAN (3G) soportan múltiples servicios: voz, datos, videollamada..., con diferente consumo de recursos [1]. Este consumo dependerá de la tasa binaria demandada. Además, unos necesitarán disponer de los recursos de forma garantizada durante la duración de la conexión, mientras que otros podrán aceptar la liberación de parte de los recursos sin que por ello se interrumpa la misma.

Llamaremos a los primeros servicios garantizados o *Circuit Switched* (CS), y a los segundos, no garantizados o *Packet Switched* (PS). Los garantizados consumen recursos de forma rígida, y los no garantizados de forma elástica.

En redes GSM (2G), el servicio de voz es del tipo garantizado (CS), y tiene un consumo rígido de recursos. Consume un circuito de 16 kbps desde la interfaz radio hasta la BSC (Base Station Controller). Este circuito discurre por redes radio y de transmisión que emplean acceso TDM.

En redes 3G, deben provisionarse recursos para una combinación de servicios CS y PS. Estos recursos ya no serán solamente circuitos TDM, sino también recursos IP.

Modelado del consumo de recursos de red

En 2G, el número de circuitos de 16 kbps necesarios para un determinado tráfico del servicio de voz (ρ Erlang), se determina mediante las tablas de Erlang B, considerando una máxima probabilidad de bloqueo admisible.

En 3G, el conjunto de servicios atendidos por el nodo consumen diferentes tipos de recursos, como por ejemplo la capacidad de procesado en banda base del nodo, la capacidad WCDMA y el ancho de banda de la interfaz Iub.

Los operadores de redes móviles deben decidir la cantidad de recursos de los que dotan a sus redes. En el caso de las redes 2G, el dimensionamiento es sencillo y puede hacerse de forma predictiva. En las redes 3G, la ausencia de modelos de dimensionamiento conduce a sobre o infradimensionamientos, provocando ineficiencias en el primer caso, o congestión en el segundo.

Detección de la degradación del Grado de Servicio (GoS)

En las redes 2G, un dimensionamiento correcto impide degradaciones en el GoS.

En el caso de las redes 3G, es habitual incurrir en degradaciones del GoS debidas a congestión - por falta de recursos -. En el caso de los servicios garantizados (como la voz), la degradación del GoS se manifiesta en forma de bloqueo, mientras que en el caso de servicios no garantizados (como los datos en modo *Best Effort*), provocará una degradación de parámetros como el *jitter*, el retardo o la tasa binaria.

Esto se podría evitar si existieran técnicas de dimensionamiento de los recursos necesarios. Al mismo tiempo, cobra importancia la afinación de técnicas de detección temprana de estas degradaciones.

1.2. Motivaciones

En la actualidad, el dimensionamiento de la capacidad en los diferentes recursos de la red UTRAN se está realizando de forma empírica. Es decir, se amplía capacidad en los diferentes recursos y se observa la evolución de los fallos de conexión en el nodo.

La presente Tesis propone un modelado de la gestión de la capacidad de los recursos de un nodo B ante la demanda simultánea de múltiples servicios. Este modelado permitirá dimensionar la capacidad de cada uno de los recursos para una combinación de tráfico multiservicio determinada, así como estimar los fallos de conexión achacables a carencias de capacidad en cada uno de los recursos. Ayudará asimismo a distinguir fallos de conexión debidos a congestión de fallos debidos a un mal funcionamiento hardware del recurso.

El resultado de este trabajo será un modelo que permita a los operadores de red conocer la distribución de recursos necesaria para atender una determinada demanda de tráfico multiservicio. Así, se podrán corregir situaciones de exceso o defecto de recursos, consiguiendo un ahorro de coste operativo para el operador y mejorando la accesibilidad de la red. Se podrán detectar situaciones de accesibilidad degradada debido a fallos hardware y se podrá realizar el dimensionamiento óptimo de la red para una demanda futura estimada. Esta aplicación es de gran interés en la actualidad, donde es habitual que en grandes concentraciones de usuarios (conciertos, manifestaciones, encuentros deportivos, ...) la red no es capaz de atender la demanda generada. Ya no tanto en voz, donde el consumo de recursos es más liviano, sino en transferencia de datos y en señalización. Los canales de tráfico ascendente suelen estar más pobremente dimensionados y en ocasiones están provocando fallos para los smartphone que intentan subir ficheros pesados a un servidor (fotos, video). En el downlink hay un uso muy intensivo (video en streaming, por ejemplo), que se ve penalizado en casos de saturación. La señalización tiene un tráfico elevado al estar estos dispositivos continuamente pasando de estado idle a estado activo (por ejemplo cuando una aplicación tipo twitter se conecta recurrentemente para comprobar si hay mensajes nuevos).

1.3. Antecedentes

1.3.1. Caracterización de los diferentes servicios

Los servicios pueden clasificarse de la siguiente forma:

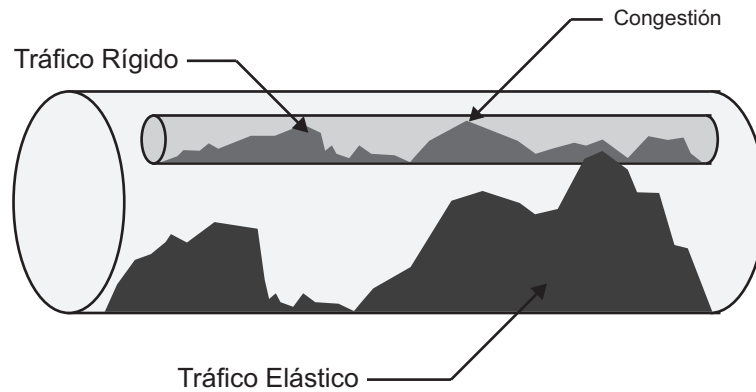


Figura 1.1: Ejemplo de tráficos garantizados y no garantizados

- ⇒ Garantizados (o rígidos): para admitir la conexión, la red comprueba si cuenta con los recursos suficientes, los reserva, y sólo los libera al finalizar la misma.
- ⇒ No garantizados (o elásticos): no se reservan recursos para una conexión, por lo que pueden ser desviados a otro servicio de mayor prioridad, reduciendo el Grado de Servicio. Adaptan su ocupación del recurso a la disponibilidad de dicho recurso de forma dinámica [2] mediante la participación de protocolos que implementan funciones tales como el control de flujo.

Un mismo servicio puede consumir de forma elástica un recurso, mientras que consume de forma rígida otro. Así, por ejemplo, el ancho de banda de plano de usuario de Iub en las conexiones de datos se consume de forma elástica, mientras que el ancho de banda de plano de control empleado en la señalización se consume de forma rígida.

Una correcta caracterización de los diferentes servicios permitirá modelar su consumo de recursos.

1.3.2. Consumo de recursos

Recursos radio

Una conexión consume fundamentalmente los siguientes recursos a nivel radio:

- ⇒ Ancho de banda determinado para el enlace entre el Nodo B y la RNC (enlace Iub). El ancho de banda disponible en cada momento dependerá de los diferentes servicios, ya que cada uno de ellos demanda una tasa binaria distinta.
- ⇒ Códigos de canalización, cuya característica más explícita son los códigos de expansión o «*spreading codes*» (SC). Este recurso es especialmente importante en el *downlink*, ya que los códigos de canalización parten de un único árbol generador, mientras que en el enlace ascendente cada usuario tiene su propio árbol de códigos. Por ello, éste será un factor limitante únicamente en el enlace descendente.
- ⇒ *Channel Elements* (CE). Se define como una unidad de medida de los recursos de procesamiento en banda base por parte de los nodos B, implementados en el *hardware*. 1 CE es igual al consumo de una conexión de voz.

- ⇒ Potencia del transmisor del nodo B. Cada tipo de tráfico soportado por la red demandará una potencia diferente, de manera que como la potencia del amplificador está acotada y no puede superar un determinado valor, da lugar a una limitación en el número de usuarios.
- ⇒ Tolerancia ante interferencia generada por el resto de conexiones del nodo. Esa tolerancia queda reflejada en la relación señal a interferencia. Este margen de calidad a priori se estrecha a medida que aumenta el número de peticiones de los diferentes servicios, hasta llegar al punto en que la calidad de la comunicación queda degradada.

Recursos de transmisión

La evolución tecnológica a nivel radio permite conexiones de mayor tasa binaria y genera mayor ocupación del enlace Iub.

Esta evolución se debe al uso de modulaciones de mayor eficiencia espectral (QPSK, 16QAM y 64QAM), al uso de doble portadora de 5 MHz (Dual Carrier) o a la transmisión múltiple (MIMO).

Por tanto, el incremento de la capacidad de la interfaz aire impone la necesidad de multiplicar la capacidad de los enlaces de transmisión en la red de transporte. Esta interfaz aire ha pasado de ofrecer unas tasas de unos pocos Mbps por sector en las primeras implementaciones a soportar máximos que en el Rel-10 del 3GPP alcanzan 168 Mbps por sector. Actualmente en el marco del Rel-11 se discute la introducción de técnicas que permitirán velocidades de pico de 672 Mbps por sector.

Del soporte exclusivo de interfaces de tipo ATM [3] se ha pasado a disponer de soporte nativo de IP [4, 5, 6, 7]. Este soporte ofrece una mucho mejor relación entre ancho de banda y coste, además de la posibilidad de formar una red de transporte común para todo tipo de servicios. En medio de esta transformación, el escenario hoy es de convivencia entre diferentes servicios y técnicas de transporte. Por un lado perviven mecanismos de envío de la información basados en la reserva rígida de cierta cantidad de ancho de banda, y por otro las técnicas de los protocolos de trama de los canales de HSPA permiten la transferencia de la información a través de la interfaz Iub con elevado factor de adaptabilidad.

Las herramientas de dimensionamiento que permiten calcular el ancho de banda necesario en una interfaz deben considerar todos los escenarios topológicos y tecnologías de transmisión y técnicas disponibles.

Algunos factores que afectan al resultado del requerimiento de ancho de banda son [4]:

- ⇒ Tecnología de transporte utilizada (PDH, SDH y Ethernet).
- ⇒ Perfil de generación de los tráficos generados por cada servicio.
- ⇒ Requerimientos de QoS para cada servicio.
- ⇒ Comportamiento de los diferentes protocolos que intervienen, en distintos niveles de la red, y que dependen entre otros del tipo de canal radio que se utilice. La acción conjunta de los distintos protocolos en los diferentes niveles de la red produce un efecto final que es preciso analizar y modelar si se quiere anticipar la calidad de servicio que experimentarán las aplicaciones en un escenario predeterminado.

1.3.3. Grado de servicio

La red reportará una serie de indicadores que permiten evaluar su GoS.

En servicios garantizados (*Circuit Switched* / CS) se utilizará la accesibilidad CS, definida como el % de conexiones que han progresado correctamente.

En servicios no garantizados se utilizarán los siguientes indicadores:

- ⇒ Accesibilidad PS (en modo paquete).
- ⇒ *Jitter* (variaciones de retardo en la llegada de paquetes).
- ⇒ Retardo o latencia de las conexiones.
- ⇒ Tasa binaria.

El bloqueo que impide que una conexión progrese correctamente tiene dos posibles causas:

- ⇒ Congestión (por ocupación de todos los recursos).
- ⇒ Hardware (cuando una avería impide ocupar un recurso disponible).

Un dimensionamiento incorrecto implicará una degradación del grado de servicio (GoS).

1.4. Objetivos

El objetivo fundamental de esta Tesis es la propuesta de una metodología de dimensionamiento para la red de acceso radio (UTRAN).

Los objetivos pueden ser clasificados de la siguiente forma:

Análisis de escenario multiservicio

- ⇒ Revisión de conceptos básicos: se analiza qué servicios ha de atender UMTS y qué recursos consume dentro de la arquitectura de la red.
- ⇒ Concreción de las diferentes aplicaciones y servicios posibles y en consecuencia su demanda de recursos para cada uno de los factores limitantes citados anteriormente.

Gestión de recursos de red

- ⇒ Estudio de las funciones de control de capacidad de la red de acceso radio basándose en el conocimiento previo de aquellas implementadas por un suministrador de red.
- ⇒ Estudio sobre una red real de las topologías de conectividad entre RBS y RNC y las estrategias de asignación de tráfico a calidades de servicio, y discusión de su adecuación y posible extrapolación del modelo a otras topologías.
- ⇒ Propuesta de una metodología de dimensionamiento de los recursos:
 - El trabajo desarrollado en esta Tesis modela la red UTRAN según un esquema multietapa, donde cada uno de los subsistemas estudia el comportamiento del tráfico frente a 3 recursos: procesado en banda base (recurso «Channel Elements»), capacidad WCDMA (recurso «Spreading Codes») y capacidad de la interfaz Iub. Este modelo determina el bloqueo producido por cada uno de los servicios, y de esta manera es posible distinguir cuál es el recurso limitante y ampliarlo.

- Detección de fallos *hardware*. El trabajo que se desarrolla en esta tesis es un modelo de congestión, y dado que se conocen los bloqueos totales del nodo mediante sus estadísticas, puede detectarse la existencia de un fallo *hardware* cuando los bloqueos por congestión que determina el modelo son menores que los bloqueos reales.
- Toda vez que se tiene un modelo de la red UTRAN, a partir de una distribución de tráfico prevista conseguir un dimensionado correcto de los recursos.

1.5. Contribuciones de la Tesis

1.5.1. Caracterización de servicios y consumo de recursos

Tal y como se verá en el capítulo siguiente, correspondiente al «estado del arte», existe una numerosa bibliografía que propone modelos para el dimensionamiento de redes multiservicio. Sin embargo, todos estos métodos se circunscriben únicamente al «ámbito académico», y su validación se realiza usualmente mediante simulaciones.

Con el desarrollo de esta Tesis Doctoral se comprobará la validez de los modelos por comparación con los valores observados en una red real. Es decir, unifica el ámbito industrial con el académico: aprovechando la ventaja que supone el conocimiento de datos estadísticos reales de tráfico cursado por estaciones de telefonía móvil y de la arquitectura real de la tecnología subyacente, será posible analizar los modelos matemáticos y demostrar su viabilidad y fiabilidad.

Asimismo, la Tesis propone correcciones al modelo teórico para ajustarlo a los patrones de tráfico reales.

1.5.2. Degradación del grado de servicio

El modelo propuesto estudia la probabilidad de bloqueo en cada un de los subsistemas que componen la red UTRAN. Posteriormente, estas probabilidades individuales se agregan para tener unas probabilidades globales para todo el sistema (tanto en sentido ascendente como descendente):

- ⇒ Accesibilidad para tráfico de circuitos (voz y videollamada) en el downlink.
- ⇒ Accesibilidad para tráfico de circuitos (voz y videollamada) en el uplink.
- ⇒ Accesibilidad para tráfico de paquetes (PS384 y HSDPA) en el downlink.
- ⇒ Accesibilidad para tráfico de paquetes (PS384 y HSUPA) en el uplink.

La contribución de este modelo radica en que en las estadísticas reales, el fabricante no proporciona este nivel de detalle. Este planteamiento que se propone en la Tesis permite detectar:

- ⇒ A partir de la equipación del nodo y del tráfico entrante, el bloqueo que provoca el dimensionado actual de los recursos, de modo que se pueda detectar cuál es el recurso limitante.
- ⇒ Combinando estas accesibilidades parciales de la manera adecuada se hace también un cálculo de la accesibilidad global para el tráfico de circuitos y de paquetes, lo que permite su comparación con los estadísticos reales para así determinar si la degradación de un nodo es debida a congestión o a un fallo *hardware*, de manera que el modelo permite detectar no solamente congestión sino también fallos de este tipo.

1.5.3. Validación de resultados en escenarios reales

La literatura que estudia la accesibilidad de sistemas UTRAN considera generalmente las limitaciones impuestas por el nivel de interferencia en el uplink o el consumo de potencia en el downlink, así como la capacidad de Iub. Y además, los modelos se prueban sobre una red simulada. Estos trabajos no se preocupan de los recursos hardware que verdaderamente limitan la capacidad del nodo, como la capacidad de procesamiento (medida en channel elements) o la capacidad de las portadoras WCDMA (en términos de spreading codes). Esta tesis analiza justamente estos recursos, junto con el consumo de Iub, y prueba el modelo desarrollado en una red real en servicio.

- ⇒ Analizando las estadísticas introduciendo como parámetros en el modelo el equipamiento del nodo y su tráfico.
- ⇒ Ampliando los recursos y comprobando su impacto en la accesibilidad total.
- ⇒ Conocidas estaciones con fallos hardware, tratando de conseguir su detección.

Gracias a esta aplicación sobre una red real, el modelo ha contribuido a:

- ⇒ Un dimensionado futuro de nodos ante eventos puntuales celebrados en las inmediaciones de la estación.
- ⇒ Un gran ahorro en inversión, ya que el operador no conocía con anterioridad el recurso degradado y ampliaba todos.
- ⇒ Mejorar la calidad de la red reduciendo la congestión.

1.6. Metodología

Propuesta de metodología de dimensionado

Para la confección y validación del modelo desarrollado a lo largo de esta Tesis, se aplica la siguiente metodología:

- ⇒ Se tiene la posibilidad de trabajar sobre una red de un operador real, teniendo acceso a estadísticas de las estaciones base o nodos B: tráfico cursado, intentos de conexión, conexiones establecidas con éxito, fallos de conexión, etc.
- ⇒ Se consideran los 3 recursos sobre los que a nivel técnico el ingeniero de red tiene acceso: capacidad de procesamiento en banda base, capacidad WCDMA y ancho de banda de la interfaz Iub entre el nodo B y la RNC. Cada uno de ellos se modela por separado para analizar el bloqueo que provoca su dimensionado sobre el tráfico de entrada. Para ello, sobre cada subsistema, se utiliza un algoritmo de cálculo de bloqueo en redes multiservicio (Kaufman-Roberts), cuyos parámetros de entrada son el tráfico y la equipación del nodo.
- ⇒ Cada cálculo del punto anterior corresponde al bloqueo asociado a cada recurso, lo que permite determinar cuál es el recurso limitante. Asimismo, las accesibilidades parciales (complementarias de los bloqueos) se combinan obteniéndose métricas comparables a los datos reportados por la red (recuérdese que el nodo no reporta accesibilidades diferenciadas por recurso, tráfico CS/PS y enlace DL/UL).

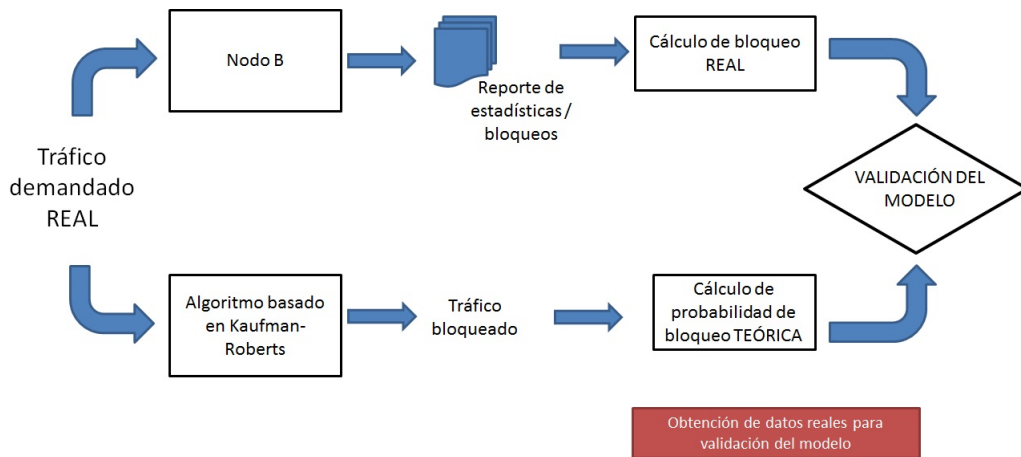


Figura 1.2: Proceso de validación

- ⇒ Se compara el grado de ocupación de recursos real que reporta la red con el estimado por el algoritmo basado en Kaufman-Roberts.
- ⇒ De la comparación del bloqueo real y la probabilidad de bloqueo teórica se determina si el nodo tiene algún recurso mal dimensionado o si la degradación es debida a un fallo hardware.

Los resultados obtenidos proporcionarán un método eficaz para un correcto dimensionamiento de los recursos de la red UTRAN.

Detección de la degradación del GoS

Por otra parte, para la propuesta del informe de estadísticas automatizado se plantea el siguiente procedimiento:

- ⇒ Se tienen datos temporales, estadísticos, de cada uno de los reportes proporcionados por los nodos de la red. Conocido el histórico de cada estación, es posible distinguir entre lo que puede considerarse un comportamiento habitual para un emplazamiento conocido, una degradación paulatina de la calidad del servicio o anomalías más concretas: incidencias puntuales o degradación persistente.
- ⇒ Posteriormente, una vez confirmada la degradación del GoS, el modelo propuesto determina si se debe a falta de recursos o a fallo hardware.

1.7. Estructura de la Tesis

La memoria de la Tesis doctoral se estructura en los siguientes capítulos:

- ⇒ El Capítulo 1 se corresponde con la introducción. En éste se incluye la motivación que ha llevado al desarrollo de esta Tesis, los objetivos que se persiguen y las principales contribuciones de la misma. Asimismo se explica la metodología que se empleará y por último, se comenta la estructura por capítulos que seguirá este documento.
- ⇒ En el Capítulo 2 se presenta de manera detallada el estado del arte. A lo largo del mismo, se hace referencia a los diferentes aspectos tratados en la Tesis y a cómo se abordan estos temas en las publicaciones más recientes.

- ⇒ El Capítulo 3 está dedicado a la arquitectura de la red de acceso radio en redes UMTS, la red UTRAN. En primer lugar se expone una perspectiva histórica acerca de la telefonía celular, haciendo un recorrido por las diferentes evoluciones que han ido aconteciendo a lo largo de las 3 generaciones. Se hace especial hincapié en los sistemas GSM y UMTS, explicando su arquitectura a nivel general junto con los elementos que la componen. A continuación se desarrolla de forma más específica cuáles son los elementos que componen la red UTRAN, concretamente las estaciones base o Nodos B, ya que en ellos se implementan varios de los recursos limitantes que se estudiarán posteriormente, como por ejemplo la capacidad de procesado en banda base a nivel hardware. Esta Tesis toma como datos estadísticas de una red real, por lo que resulta conveniente, y así se ha hecho, la particularización de la arquitectura general según el fabricante Ericsson. Asimismo, se hace también una introducción a la tecnología LTE (*Long Term Evolution*), su red de acceso radio E-UTRAN y en última instancia se realiza una comparativa entre UTRAN y E-UTRAN. La última parte del capítulo está centrada en las interfaces que están definidas entre las distintas entidades de la red UMTS. Este desarrollo resulta de especial importancia, ya que la implementación de la interfaz entre el Nodo B y la RNC (interfaz Iub) es otro factor limitante objeto de estudio en la presente Tesis. A lo largo de esta parte se hace una distinción a nivel de capas según el modelo OSI, contemplando en primer lugar las distintas tecnologías implementadas en la capa física (radioenlaces, xDSL, fibra óptica...), y en segundo lugar el nivel de enlace (tecnología ATM y, muy importante, MPLS por sus mecanismos que dotan la capacidad de priorizar tráfico). Al igual que en el capítulo anterior, procede particularizar la estructura general sobre la red de trabajo de Ericsson. El último apartado de este capítulo está dedicado a las interfaces en redes LTE.
- ⇒ El Capítulo 4 presenta el algoritmo mediante el cuál se implementa el cálculo de la accesibilidad. Así, se describe en primer lugar la situación bajo la cual se realizarán los cálculos posteriores, justificando el perfil como una red multiservicio, en el que distintos servicios demandan diferente cantidad de un mismo recurso compartido por todos ellos y por el cual compiten. Se presenta una primera alternativa basada en el algoritmo original, y cómo el autor de la Tesis ha solventado las limitaciones que éste presenta para sus cálculos particulares. Así, diferentes aspectos de su implementación han dado lugar a distintas modificaciones del mismo para permitir los cálculos, pasando de un algoritmo recursivo a otro completamente distinto basado en la FFT. Finalmente, se concluye el capítulo con una comparativa de rendimiento de todos los algoritmos estudiados.
- ⇒ El Capítulo 5 tiene como objeto la caracterización de los diferentes servicios que soportará la red. Se define el concepto de servicios portadores y se detallan las distintas clases de tráfico definidas por el estándar de UMTS, comparándolas todas ellas y presentando los diferentes parámetros de calidad de servicio asociados. Finalmente, se detallan los atributos de calidad de servicio y las diferentes funciones de gestión de QoS a proporcionar en una red UMTS, así como su posible ubicación.
- ⇒ En el Capítulo 6 se describen los recursos disponibles en la red, así como sus roles limitantes, que serán objeto de estudio a lo largo de la presente Tesis. Se ubica al lector en una metodología de trabajo consistente en analizar cada uno de los recursos como subsistemas independientes cuya contribución da lugar a la accesibilidad total. Así, apoyándose en la

descripción particular comentada a lo largo de los capítulos 2 y 3, se describen ampliamente estos recursos.

- ⇒ En el Capítulo 7 se particulariza el modelo de Kaufman-Roberts para la red de acceso UTRAN. Se ubica al lector en una metodología de trabajo consistente en analizar la accesibilidad total del sistema como un producto de accesibilidades parciales derivadas del empleo de cada uno de los recursos descritos en el capítulo 6. De esta manera, se pretende dar una visión del conjunto como una sucesión de subsistemas, cada uno de los cuales puede entonces analizarse y modelarse por separado. En las siguientes páginas se particulariza el algoritmo multiservicio presentado en el Capítulo 4 a cada uno de los subsistemas, particularizando para cada recurso los parámetros de entrada y salida. Se hace especial hincapié en cuáles son los recursos compartidos por los que compiten los diferentes tráficos y cómo se van consumiendo los mismos a medida que aumenta la demanda del servicio. Como continuación, se detalla la implementación del algoritmo de Kaufman-Roberts. El estudio de la accesibilidad cobra especial importancia en situaciones en las que el tráfico se incrementa de forma repentina. Teniendo presente que las estadísticas reportadas por el nodo tienen una resolución temporal claramente insuficiente por el contexto de estudio de la Tesis, se presentan algunas soluciones de interpolación basadas en distribuciones de probabilidad. Así, el conformado del tráfico cobra una importancia crítica y, conocida la accesibilidad real por las estadísticas del nodo, es posible calibrar el sistema. A continuación, el capítulo describe las diferentes alternativas que el autor de esta Tesis ha considerado a la hora de realizar una implementación teórica, justificando la elección de Matlab como herramienta de trabajo.
- ⇒ El Capítulo 8, derivado del anterior, procede a analizar comparativamente los resultados del modelado teórico con los datos reales de que se dispone gracias a los reportes obtenidos por estaciones de una red real. Como ya se ha comentado, una de las contribuciones de esta Tesis consiste en aunar por una parte un escenario teórico analizado en un ámbito académico con datos reales de una red UMTS con aplicación en un ámbito más industrial.
- ⇒ En el Capítulo 9 el autor expone las conclusiones de la labor que ha desarrollado a lo largo de la presente Tesis, así como las posibles futuras vías de investigación que pueden derivarse de este trabajo.
- ⇒ En el Apéndice A se anexa el código fuente implementado en Matlab, descrito a lo largo de los Capítulos 6 al 7.
- ⇒ En los Apéndices B y C se presentan con mayor profundidad algunos conceptos teóricos ampliando la información correspondiente al capítulo 3, dedicado a las interfaces UTRAN y, en particular, a las tecnologías de capa física ATM y PDH/SDH.
- ⇒ En el Apéndice D se describe el procedimiento de cálculo del ancho de banda medio utilizado en el RAB de Iub sobre ATM.
- ⇒ En el Apéndice E se amplía la información del Capítulo 2 desde el punto de vista de protocolos de las interfaces de la red UTRAN, así como los diferentes mecanismos de gestión de recursos radio.
- ⇒ En el Apéndice F se presenta una descripción de los KPIs de una red Ericsson. Dado que la Tesis toma como base estadísticos reportados por una red real Ericsson, resulta pertinente

enumerar cuáles son los KPIs (*Key Performance Indicators*) contemplados, junto con otra información relevante relacionada con el fabricante.

- ⇒ En el Apéndice G el autor presenta cómo las técnicas de aprendizaje y de inteligencia colectiva pueden servir de aplicación en el contexto en el que se desarrolla la Tesis.
- ⇒ En el Apéndice H se detalla el material empleado para la elaboración de esta Tesis, en cuanto a herramientas de trabajo para la elaboración de documentos como de herramientas de cálculo para la implementación del estudio teórico seleccionadas en el capítulo 7.
- ⇒ Finalmente, se presenta un glosario con los acrónimos empleados a lo largo de la Tesis, la bibliografía seleccionada como punto de partida para el estudio del estado del arte, así como todas las obras a las que se hace referencia a lo largo de este documento.

Capítulo 2

Estado del arte

2.1. Estado del arte

En este capítulo se describen los diferentes estudios que se han realizado recientemente en el ámbito del análisis de la capacidad de redes UTRAN. Puede destacarse que un común denominador de estas referencias es el análisis del consumo de recursos validado mediante simulaciones de la red. El modelo que se desarrolla en esta Tesis ha sido validado con datos reportados por datos de una red real en servicio; este hecho ha motivado que este estudio no alcance el modelado de redes LTE, al no disponer en España en 2012 de nodos LTE en servicio (sólo algunos cluster de pruebas en Madrid y Barcelona).

Por otra parte, en la mayoría de la literatura se estudia el impacto del diseño de la red en su capacidad. Así, por ejemplo, se establecen recomendaciones en términos de optimización de la calidad de la interfaz radio, reduciendo sobrealcances y minimizando el solape entre nodos. La reducción de sobrealcances permite evitar que un Scrambling Code de un nodo reduzca la E_c/N_0 en el ámbito de cobertura de nodos lejanos, ya que su señal allí será ruido (al no formar parte del Active Set). Además, la reducción del área de soft handover redundante en un decremento del número de celdas en el Active Set, bajando así la E_c/N_0 debida a interferencia entre nodos, y reduciendo el factor de soft handover (lo que hace que dos nodos transmitan hacia la RNC por duplicado el flujo de datos correspondiente a la conexión de un usuario).

El trabajo llevado a cabo en esta Tesis, por el contrario, se centra en el estudio de la ocupación en aquellos recursos que son susceptibles de ampliación, como son las tarjetas de procesado en banda base, la capacidad Iub y el número de portadoras WCDMA (con sus árboles asociados). Este estudio es eminentemente aplicado y cubre una necesidad clara en los operadores de redes móviles, ya que hasta el momento no cuentan con herramientas de dimensionado para redes 3G. La falta de estas herramientas conduce a dimensionado ineficiente (teniendo recursos infrautilizados por ampliaciones no necesarias, congestión no detectada en determinados recursos, incapacidad de anticipación a picos de demanda de capacidad, etc). El coste económico de no disponer de este tipo de herramientas es muy significativo, lo que ha motivado el estudio.

A continuación, se describen las principales referencias analizadas en la revisión del estado del arte. Las tareas de operación y mantenimiento de redes representan un alto coste para los operadores de red, ya que requieren de mano de obra cualificada. La automatización en la detección de fallos de red ya ha sido tratada en otros trabajos. En [8] se presenta una metodología de automatización para la resolución de incidencias en la red utilizando una aproximación bayesiana.

En su trabajo los autores analizan una relación entre la sintomatología y fallos conocidos en la red, lo que les permite construir un modelo estadístico para identificar futuros fallos, basados en las observaciones. El modelo está basado en el simulador de redes UMTS llamado Odyssee, desarrollado por France Telecom. Por el contrario, esta tesis aprovecha datos reales tras haber analizado tareas de operación de red para solucionar problemas de capacidad.

La optimización del diseño de la red también se ha estudiado de manera extensa. En [75] se presentan estrategias para el ajuste de parámetros en antenas y de potencia transmitida, basado en sistemas de aprendizaje. Otros estudios se han enfocado en modelar el tráfico para el dimensionamiento de la interfaz en WCDMA [76, 77, 78, 9] o el Iub [79, 80, 81, 82]. El trabajo presentado en esta tesis utiliza el modelo calibrado con datos reales de la red para una identificación efectiva del recurso congestionado, de manera que los operadores puedan distribuir adecuadamente los recursos de capacidad en la red.

El trabajo realizado representa con exactitud la accesibilidad desglosada por recurso. Por tanto, puede utilizarse fácilmente para predecir el comportamiento del sistema e identificar el recurso limitante para demandas de tráfico previstas. Puede así determinarse la capacidad necesaria para eventos futuros.

Una serie de autores analizan el impacto del nivel de interferencia uplink en la capacidad del nodo. Como es sabido, el factor de carga generado por la demanda de servicios de los usuarios atendidos por el nodo, junto con el ruido que recibe el nodo de conexiones de equipos de usuario (Ues) con otros nodos, elevan el factor de interferencia. Antes de llegar al punto de desensibilización, el nodo descartaría conexiones para reducir el factor de carga. Así, de acuerdo con [10], el bloqueo se produce cuando la interferencia total excede el umbral. En el uplink la capacidad está limitada por la interferencia total producida por la carga de usuarios y el nivel de ruido en el uplink debido a señal de otras celdas. Proponen un algoritmo tipo Kaufman-Roberts modificado para operar con su modelo.

En otro trabajo [11], se obtiene un modelo para la capacidad del nodo considerando como recurso limitado la potencia downlink del nodo.

Hay otro conjunto de autores que implementan versiones modificadas del algoritmo de Kaufman-Roberts. En la presente Tesis se utiliza la versión general, pero en su implementación FFT, en lugar de la recursiva (para reducir la complejidad computacional).

En [12, 13] se propone una extensión de la fórmula de Kaufman-Roberts para ser aplicada en servicios elásticos de CDMA, donde el cálculo de la probabilidad de bloqueo no es trivial. Se utiliza el modelo de [14] para el uplink.

Para la presente Tesis se han tomado como inspiración trabajos como [15] para proponer el método de agregación de accesibilidades en los distintos recursos del sistema. También [90] ha sido de ayuda a la hora de planificar el estudio del bloqueo multiservicio en sistemas multi-recurso.

La referencia [16] ha servido para comprender cómo se asignan los recursos a nivel CS y PS dentro de la interfaz Iub. También han sido de utilidad [79, 80, 17] para determinar la ocupación de Iub aportada por HSDPA y HSUPA. [18] es útil para dimensionar la capacidad del Iub.

Artículos como [10] analizan el impacto de la demanda de los diferentes servicios usando el mismo enfoque que la presente Tesis (Kaufman Roberts), pero sobre un único recurso: el nivel de interferencia UL. En la Tesis se aplicará sobre recursos en los que el operador de red tiene capacidad de actuación.

El texto [19] sirve de base para proponer el uso de sistemas de aprendizaje supervisado para confirmar si las predicciones de este modelo son correctas, corrigiéndolo si fuese oportuno.

Algunos autores se centran en mejorar el diseño de la red 3G para optimizar su rendimiento, y mejorar las condiciones de la interfaz radio, como [20] o [21]. Para la Tesis, el autor parte de una red ya diseñada y en servicio, a la que hay que dotar de la capacidad necesaria.

Algunos autores como [22] o [23] proponen algoritmos que optimizan la ocupación de los recursos radio; no es objeto de este trabajo proponer algoritmos diferentes a los estandarizados por 3GPP, sino comprender cómo ocupan los recursos para poder anticiparse a su bloqueo.

En la sección en la que se describe la demanda de los diferentes servicios ha servido de inspiración [24] o [78].

Para comprender los Key Performance Indicators (KPIs) que debían usarse como referencia han servido de base la experiencia del operador y textos como [83].

Capítulo 3

Arquitectura de red UTRAN

3.1. Introducción

Durante las dos últimas décadas, el mercado de las telecomunicaciones ha asistido al desarrollo de dos tecnologías que han supuesto una absoluta revolución social y económica; por un lado, Internet ha comunicado el mundo, y por otro, las redes móviles (figura 3.1) permiten la comunicación desde cualquier punto. Este documento se centra en el estudio de una de las partes que componen las redes de comunicación «3.5G» actuales, aunque previamente se realizará un repaso de su evolución desde los inicios de las comunicaciones móviles hasta llegar a las últimas redes comerciales, que dejan entrever ya lo que serán las redes «4G» que llegarán en los próximos años.

3.1.1. Primeros Pasos

Elegir un momento histórico a partir del cual contar la evolución de las comunicaciones móviles hasta nuestros días no es fácil ya que muchos han sido los individuos, organizaciones, universidades y empresas que han contribuido a este desarrollo. Tal vez el punto de partida que más partidarios tenga se remonte al año 1831, cuando Michael Faraday demostró el principio de inducción electromagnética. Ese mismo año nacía en Edimburgo James Clerk Maxwell quien, en 1873, publicaría su obra *Treatise on Electricity and Magnetism* donde se recogían las bases teóricas de los campos

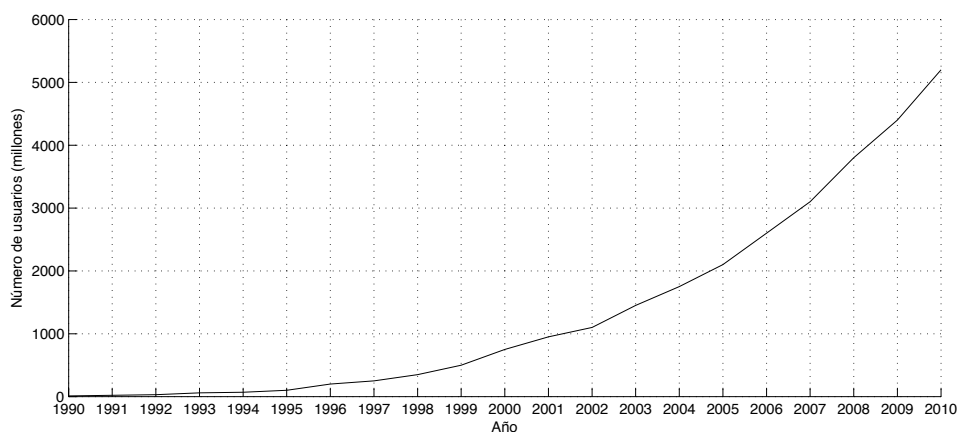


Figura 3.1: Evolución del nº de usuarios de telefonía móvil en los últimos 20 años

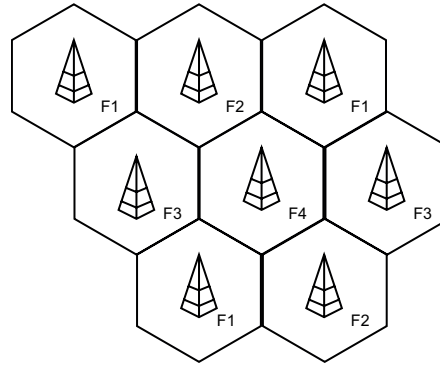


Figura 3.2: Concepto de red celular

electromagnéticos y se enunciaban sus famosas ecuaciones. Tras Maxwell y Faraday, Heinrich Hertz conseguiría desarrollar el primer oscilador y antena capaces de transmitir y detectar ondas.

Tras Hertz, las investigaciones puramente teóricas comienzan a llevarse al mundo de la ingeniería, y tanto Nikola Tesla como Guglielmo Marconi comienzan a realizar las primeras demostraciones de comunicaciones inalámbricas. Es a partir de este momento cuando lo que hasta entonces habían sido meros experimentos científicos se convierten en medios de comunicación comerciales: en el año 1906 nace la primera radio y al año siguiente se establece la primera línea comercial transoceánica. La evolución es imparable. En 1927 se produce la primera transmisión de televisión entre América y Europa, en 1933 se inventa la modulación FM, se inventan los primeros transmisores/receptores móviles que se equipan en los coches de policía, etc.

3.1.2. Telefonía Celular

A pesar de que las comunicaciones móviles ya se comienzan a realizar en los años 30 con equipos a bordo de los coches policía, estos sistemas eran muy limitados: su radio de cobertura era de pocos kilómetros, los equipos eran grandes y pesados y tenían una capacidad máxima de usuarios baja (de 50 a 80 abonados). Durante las siguientes décadas estas redes fueron evolucionando: se las dotó de más alcance, más capacidad y con dispositivos menos voluminosos. Sin embargo, si se quería crear una red inalámbrica con una cobertura nacional y capaz de ser usada por miles de abonados era necesario desarrollar una nueva tecnología. Ésta nace a partir del concepto de red «celular» presentado por los laboratorios Bell en el año 1947, pero que no será desarrollada hasta la década de los años 70. Este concepto se reduce a dividir a la red en pequeñas áreas denominadas celdas. Cada una de ellas dispone de una estación base receptora/transmisora, con la que se comunican los abonados, enlazada con la estación de conmutación telefónica de la red fija. Cada una de estas celdas opera normalmente con su propia frecuencia, a fin de evitar interferencias con sus celdas vecinas; sin embargo, las situadas lo suficientemente lejos para no interferirse emplearán la misma frecuencia, ya que se dispone de un número limitado de éstas.

Gracias al uso de celdas se conseguía solucionar los problemas de cobertura y capacidad de los sistemas anteriores. Simplemente añadiendo nuevas celdas la red podía crecer en superficie de cobertura y, reduciendo el área de las mismas y aumentando el número, se podía dotar de más capacidad a la red.

3.1.2.1. Primera Generación

Bajo el concepto celular, a comienzos de los años 80 comienzan a surgir las primeras redes comerciales: surgen lo que posteriormente se ha denominado redes de primera generación. Son redes de carácter regional y nacional, por tanto, incompatibles entre sí. Además son redes analógicas y que nacen con el espíritu de ampliar las redes fijas existentes. Dentro de estas redes nos encontramos con el sistema *Nordic Mobile Telephone* (NMT) y el *American Phone System* (AMPS). Son redes que ofrecían servicios básicos de voz.

Su éxito fue mayor al esperado: la necesidad de disponer de redes móviles globales fue percibida a mitad de los años 80, lo que lleva a las instituciones de estandarización internacionales a darse cuenta de que es necesario definir un nuevo sistema más ambicioso.

3.1.2.2. Segunda Generación

Posiblemente el gran éxito de la telefonía móvil sea debido al propio planteamiento a mediados de los años 80 de lo que debía ser la segunda generación. Se fijó como principal objetivo el desarrollo de una red global, totalmente compatible y bajo un estándar que definiese exhaustivamente todos los aspectos necesarios para que los fabricantes de los equipos desarrollasen equipos que pudieran entenderse entre sí. Además se pensó que estas redes no debían limitarse en exclusiva a ofrecer servicios de voz: se pensó en diseñar una arquitectura capaz de crecer en servicios.

De los objetivos iniciales nacieron tres redes: IS-54/IS-94 en Estados Unidos, JDC en Japón y GSM en Europa. De estas tres fue GSM, siglas de *Global System for Mobile Communications*, la red que más éxito tuvo. De hecho no fue implantada sólo en Europa, sino que su crecimiento llegó a Asia, África y parte de América y se convirtió en la referencia global de la telefonía móvil.

GSM

GSM (*Global System for Mobile Communications*, aunque originalmente significaba *Groupe Spécial Mobile*) se convirtió en el estándar de telefonía móvil más popular a lo largo del mundo. Se estima que casi el 80 % del mercado de redes móviles opera bajo este sistema, lo que se traduce en más de 1200 millones de usuarios en 212 países.

GSM fue presentada por primera vez en 1991. Fue el primer sistema en especificar una red móvil completamente digital estructurada en distintos niveles y servicios. Inicialmente se desplegó utilizando frecuencias comprendidas entre 890-915 MHz para el enlace ascendente (*Uplink* ó UL) y 935-960 MHz para el enlace descendente (*Downlink* ó DL); a este sistema se le conoce como GSM900. Posteriormente, se desarrolla otra versión funcionando en bandas de 1800 MHz denominada DCS (*Digital Cellular System*).

La arquitectura de GSM, mostrada en la figura 3.3, fue copiada por distintas tecnologías, entre ellas la propia UMTS que no es más que una evolución. Se definen tres subsistemas independientes:

- ⇒ Mobile Station (MS): comprende a la parte de usuario y en ella se encuentra definidas las características del equipo hardware y de lo que se conoce como tarjeta SIM.
- ⇒ Base Station Subsystem (BSS): es la parte responsable de la gestión del tráfico y la señalización asociada entre los móviles y el núcleo de red o Network Switching Subsystem. La BSS se encarga de toda la gestión de la parte radio y movilidad de los usuarios.

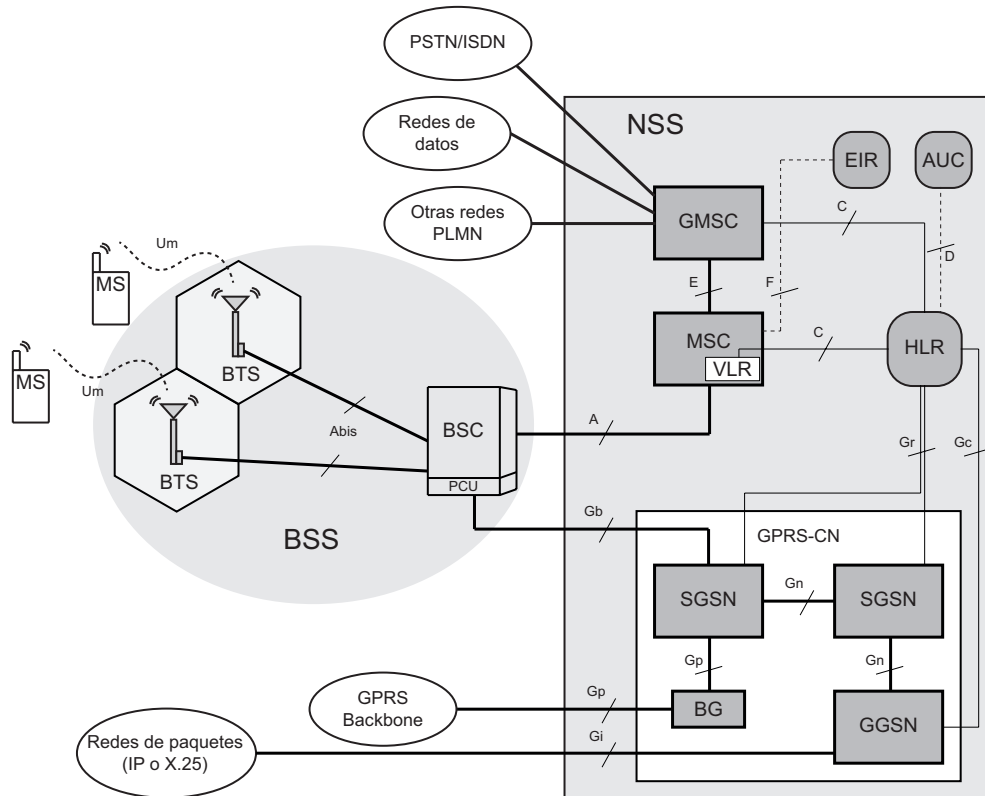


Figura 3.3: Esquema de la arquitectura de una red GSM

⇒ Network and Switching Subsystem (NSS), es el componente que realiza las funciones de portar y administrar las comunicaciones entre teléfonos móviles y la Red Conmutada de Telefonía Pública (PSTN).

3.1.2.3. Tercera Generación

La historia de las comunicaciones móviles de tercera generación comienza a gestarse a mediados de los años 80 en el seno de la *International Telecommunication Union* (ITU). Durante los años 90 la ITU comienza a definir un conjunto de requerimientos para las redes 3G, llamado IMT-2000. Por aquel entonces parecía imposible conseguir un único estándar mundial a tenor de lo que estaba sucediendo en las redes 2G del momento. Cada fabricante, operador o país trataba de imponer su tecnología y deseaba que la siguiente evolución supusiera el mínimo impacto en sus respectivas redes 2G. Así pues, IMT-2000 se convirtió en un paraguas para albergar las diferentes evoluciones de 2G: UMTS, CDMA2000 y WiMaX. Todos los miembros del IMT-2000 debían satisfacer los mismos niveles de calidad, disponer de una estructura similar y permitir la interoperabilidad entre ellas. También a mediados de los 90, la Unión Europea promovió un proyecto común de estandarización de una nueva tecnología móvil: comenzó la estandarización de la red UMTS que implicó a un gran número de fabricantes y operadores.

A nivel técnico, puede decirse que muchos de los aspectos de UMTS son en realidad requerimientos del IMT-2000. En primer lugar existe el concepto «*Anytime anywhere communication*», que podemos resumir en:

⇒ Red compartiendo el mismo estándar.

- ⇒ Espectro unificado, es decir, la misma banda de frecuencias reservada a nivel mundial (GSM, por ejemplo, no cumplía este requisito).
- ⇒ Por último, alta cobertura. Este punto fue promovido por las distintas autoridades reguladoras.

En segundo lugar, se debía dotar a la red de capacidad de transporte de datos con capacidades de acceso a Internet y servicios multimedia. Esto suponía tener que ofrecer una gran cantidad de ancho de banda. IMT-2000 fijó una serie de objetivos al respecto:

- ⇒ Capacidad de descarga de hasta 2 Mb/s para entornos fijos y 144 kb/s para situaciones de movilidad.
- ⇒ Soporte para tasas de bit simétricas y asimétricas.
- ⇒ Soporte para servicios con tasas de bit variable.
- ⇒ La calidad de la voz debía ser, al menos, tan buena como lo era en las redes 2G.
- ⇒ Se debía dejar la posibilidad para añadir nuevos servicios en el futuro que permitiesen crecer al estándar.

Además también se fijaron una serie de requisitos para los operadores:

- ⇒ La red debía ser altamente segura, incluyendo funcionalidades de autenticación y autorización.
- ⇒ Se debía implementar una funcionalidad de tarificación más sofisticada que las existentes.
- ⇒ La red debía ser retrocompatible con las existentes redes 2G.

Bajo todas estas premisas el grupo de trabajo 3GPP lanzó en el año 1999 la primera versión de UMTS conocida como **Release 99 (R99)**. Esta **release** especificaba al completo el funcionamiento de UMTS y cumplía los objetivos establecidos por el IMT-2000. Sin embargo, desde el primer momento no se vio a esta especificación como definitiva, sino como un punto de partida sobre el que trabajar para seguir evolucionando las tecnologías de tercera generación. De hecho, en los más de diez años que han pasado desde el nacimiento de R99 el foro 3GPP ha redactado otras 10 **releases** que se han introducido y han mejorado funcionalidades de la original: de los 2 Mbps originales se ha evolucionado a velocidades de hasta 84 Mbps gracias al desarrollo de HSPA+ basado en el uso de modulaciones más altas, nuevos canales de tráfico y técnicas MIMO. Se espera que la próxima **release 11** establezca un nuevo récord de 672 Mb/s gracias al uso de multiportadoras.

Cuando los operadores de telefonía móvil comenzaron el despliegue de redes UMTS la mayoría de ellos disponían de redes GSM previas. Antes de analizar en profundidad las redes UMTS, en el siguiente apartado se describe sucintamente cómo se estructuran las redes GSM, punto de partida de la tecnología 3G.

3.1.3. Arquitectura de redes GSM

En la figura 3.3 se ilustra la red de referencia GSM, donde se muestran tanto las interfaces como los elementos necesarios para soportar tanto los dominios CS (*Circuit Switched*, circuitos conmutados) y PS (*Packet Switched*, paquetes conmutados). En esta red de referencia se pueden distinguir tres subredes:

- ⇒ *Base Station Sub-System* (BSS) o *GSM/Edge Radio Access Network* (GERAN). Este subsistema está principalmente compuesto de la *Base Transceiver Station* (BTS) y la *Base Station Controller* (BSC), que entre ambos controlan las interfaces de radio GSM. Con la incorporación de las funcionalidades de transferencia de datos a GSM con el desarrollo de *General Packet Radio Service* (GPRS), se añadió un nodo adicional entre el GPRS-CN y la interfaz radio, esto es el *Packet Control Unit* (PCU).
- ⇒ *Network and Switching Sub-System* (NSS). Principalmente constituido por la *Mobile Switching Center* (MSC), encargada de enrutar las llamadas a los terminales móviles. La MSC necesita de una serie de bases de datos y registros (VLR, HLR, EIR y AuC) para diversas tareas administrativas y de control. Como punto de interconexión hacia otras redes se tiene el *Gateway MSC* (GMSC), que en la práctica se encuentra siempre formando parte del mismo equipo que la MSC.
- ⇒ *General Packet Radio Service, Core Network* (GPRS-CN). Situado dentro del NSS y formado por otros dos nodos el *Serving GPRS Support Node* (SGSN) y el *Gateway GPRS Support Node* (GGSN). Ambos guardan un paralelismo con los nodos MSC y GMSC pero en el dominio de circuitos conmutados.

Todos los elementos de la red GSM se comunican entre sí mediante interfaces y protocolos definidos en el estándar y, como veremos a continuación, constituyen la base fundamental de la primera versión de UMTS.

3.2. Redes UMTS. Arquitectura de red de acceso radio

En este apartado se presenta una visión detallada de la arquitectura de UMTS, incluyendo una primera introducción a los elementos lógicos que componen la red y a sus interfaces. UMTS utiliza la misma arquitectura, ya conocida, que la que han empleado los principales sistemas de segunda generación, e incluso algún sistema de primera generación. UMTS consiste en un conjunto de elementos de red lógicos, cada uno de los cuales tiene definida una funcionalidad. En los estándares, los elementos de red se definen a nivel lógico, pero es muy común que esa definición se lleve después al nivel físico, especialmente dado que hay un conjunto de interfaces «abiertas» (para que una interfaz se denomine «abierta», el requisito es que haya sido definida con tanto nivel de detalle que los equipos en los extremos puedan ser de dos fabricantes diferentes). Los elementos de la red se pueden agrupar basándose en funcionalidades similares o basados en a qué sub-red pertenecen.

Funcionalmente, los elementos de red se agrupan en la red de acceso radio (RAN; *UMTS Terrestrial RAN (UTRAN)*) que se encarga de toda la funcionalidad relacionada con la radio, y el núcleo de red (CN), que es responsable de la conmutación y el enrutado de llamadas y conexiones de datos hacia redes externas. Para completar el sistema, se define también el terminal o equipo de usuario (UE), límite entre el usuario y la red radio. Un esquema de la arquitectura del sistema a grandes rasgos se muestra en la figura 3.4.

Desde el punto de vista de las especificaciones y los estándares, el UE y la red UTRAN se han definido con protocolos completamente nuevos, cuyos diseños están basados en las necesidades de la por aquel entonces nueva tecnología radio WCDMA. Por el contrario, la definición del núcleo de red se adoptó la de GSM. Este hecho proporciona a un sistema con una nueva tecnología radio un núcleo de red muy desarrollado y robusto, que acelera y facilita su introducción y proporciona ventajas muy competitivas como el roaming global.

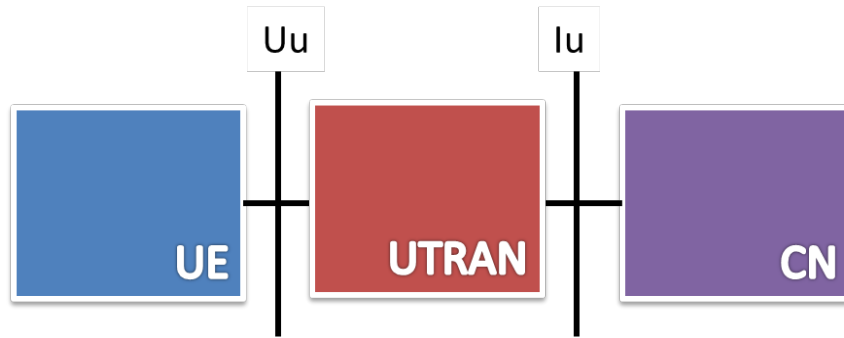


Figura 3.4: Arquitectura general de UMTS

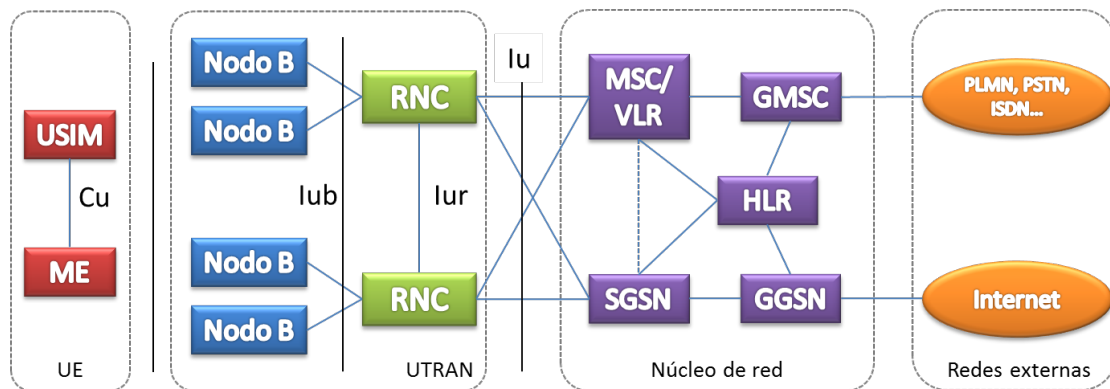


Figura 3.5: Elementos de la red en una PLMN

Otra manera de agrupar los elementos de la red UMTS es dividirlos en sub-redes. UMTS se puede considerar modular en el sentido de que es posible tener varios elementos de red del mismo tipo. En principio, el mínimo requisito para tener una red completa y operativa es tener al menos un elemento lógico de cada tipo (algunas características, y por tanto algunos elementos de la red, son opcionales). La posibilidad de tener varias entidades del mismo tiempo permite la división de UMTS en diferentes sub-redes que son operativas en sí mismas o juntas con otras sub-redes, y que se distinguen unas de otras con identificadores propios. A cada sub red se la conoce como UMTS PLMN (UMTS Public Land Mobile Network). Típicamente, una PLMN es manejada por un único operador y se conecta a otras PLMN como si fuera cualquier tipo de red, como RDSI, PSTN, Internet, etc. La figura 3.5 muestra los elementos en una PLMN (para poder ilustrar también las conexiones, incluye redes externas).

La arquitectura de la red UTRAN se desarrollará con detalle en la sección siguiente, si bien a continuación se presenta una pequeña introducción a todos sus elementos.

El terminal o equipo de usuario (UE) está formado por dos partes:

- ⇒ El equipo (ME - *Mobile Equipment*), es el terminal radio utilizado para la comunicación radio sobre la interfaz Uu.
- ⇒ El USIM (*UMTS Subscriber Identity Module*) es una tarjeta inteligente que recoge la identidad del cliente, implementa algoritmos de autenticación y almacena claves de encriptación e información sobre el cliente que será necesaria para el terminal.

La UTRAN también está formada por dos elementos:

- ⇒ El nodo B, convierte el flujo de datos entre las interfaces Iub y Uu. También participa en la gestión de recursos radio.
- ⇒ La controladora de red radio (RNC - *Radio Network Controller*) controla los recursos radio en su dominio (el nodo B está conectado a ella). La RNC es el servicio de punto de acceso para todos los servicios que proporciona la UTRAN al núcleo, como la gestión de las conexiones al UE.

Los principales elementos del núcleo de red GSM (hay otras entidades no mostradas en la figura), son las siguientes:

- ⇒ Registro HLR (*Home Location Register*) es una base de datos localizado en el sistema local del usuario que almacena la copia principal del perfil del usuario. Este perfil contiene por ejemplo información sobre los servicios permitidos, áreas de roaming prohibidas e información sobre servicios suplementarios como el estado de un desvío de llamadas, etc. Se crea cuando un nuevo usuario se suscribe al sistema, y permanece almacenado tanto tiempo como el usuario permanezca siendo cliente. Para el propósito de encaminar llamadas o mensajes cortos, el HLR también almacena la localización del UE a nivel de MSC/VLR y/o SGSN.
- ⇒ Registro MSC/VLR (*Mobile Services Switching Centre/Visitor Location Register*) es el conmutador MSC y base de datos VLR que sirve al UE en su localización actual para los servicios de conmutación de circuitos (CS). La función del MSC aquí es conmutar las peticiones de CS y la función del VLR guardar una copia del perfil del usuario visitante, así como información más precisa sobre la localización del UE dentro del sistema. La parte de la red a la que se accede vía el MSC/VLR se la conoce a menudo como «dominio de circuitos», o «dominio CS».
- ⇒ *Gateway MSC* (GSMC). Es el conmutador en el punto en el que UMTS PLMN se conecta a redes de circuitos externas. Todas las conexiones entrantes y salientes pasan a través del GSMC.
- ⇒ SGSN (*Serving General Packet Radio Service (GPRS) Support Node*). Su función es similar a la del MSC/VLR pero se usa normalmente para los servicios de conmutación de paquetes (PS).

Las redes externas pueden dividirse en dos grupos:

- ⇒ Redes de circuitos (*CS Networks*). Están provistas de conexiones de conmutación de circuitos, como el servicio telefónico tradicional. Otros ejemplos de redes CS son el RDSI y la red telefónica pública conmutada (PSTN).
- ⇒ Redes de paquetes (*PS Networks*). Están provistas de conexiones para paquetes de datos. Internet es un ejemplo de red PS.

Los estándares de UMTS están estructurados de tal manera que la funcionalidad interna de los elementos de la red no se especifica con mucho detalle. En su lugar, las que están sí están definidas son las interfaces entre los elementos lógicos. Están especificadas las siguientes:

- ⇒ Interfaz Cu: es la interfaz eléctrica entre la tarjeta USIM y el ME. Sigue un formato estándar para tarjetas inteligentes.

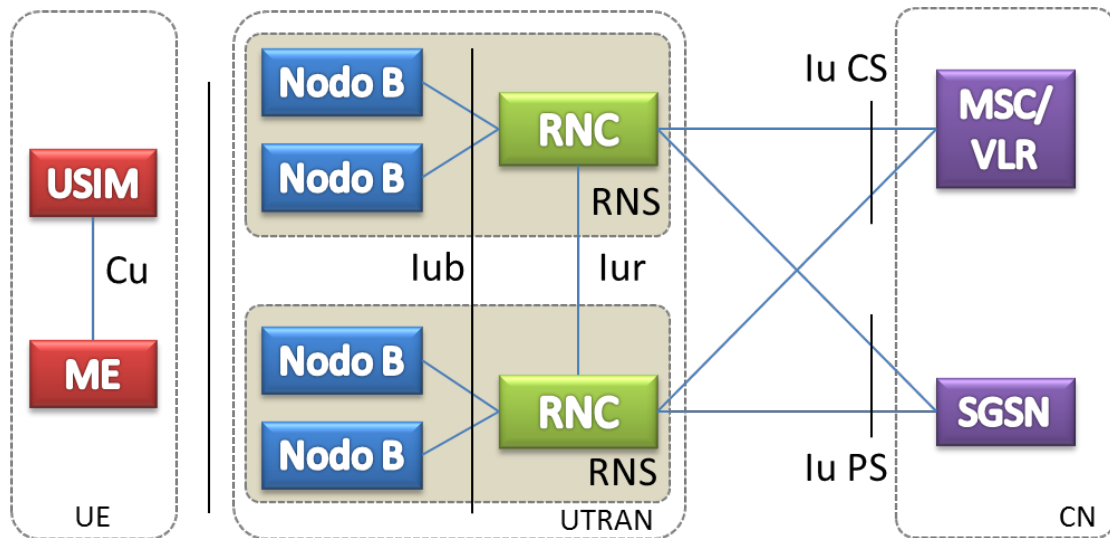


Figura 3.6: Arquitectura de red UTRAN

- ⇒ Interfaz Uu: es la interfaz radio de WCDMA. La interfaz Uu es a través de la cual el UE accede a la parte fija del sistema y, por tanto, es probablemente la interfaz abierta más importante de UMTS. Indudablemente hay más fabricantes de equipos de usuario que de elementos fijos de red.
- ⇒ Interfaz Iu: conecta la red UTRAN con el núcleo de red. Similar a las interfaces correspondientes de GSM, A (CS) y Gb (PS), la interfaz abierta Iu proporciona a los operadores de UMTS la posibilidad de adquirir equipos de UTRAN y CN de diferentes fabricantes.
- ⇒ Interfaz Iur: es abierta y permite soft handover entre RNCs de diferentes fabricantes y, por tanto, complementa la interfaz Iu.
- ⇒ Interfaz Iub: conecta un nodo B con la RNC. UMTS es el primer sistema de telecomunicaciones móviles comercial donde la interfaz de la controladora de estaciones base está estandarizada como una interfaz libre.

3.3. Arquitectura de la red UTRAN

La arquitectura general de la red UTRAN está representada en la figura 3.6. Está compuesta de uno o varios subsistemas RNS, los cuales son subredes dentro de la red UTRAN y están compuestos de una RNC y uno o varios Nodos B. Las RNCs pueden estar conectadas unas a otras a través de la interfaz Iur. Las RNCs y los nodos B están conectados mediante al interfaz Iub.

Antes de entrar en una descripción más detallada de los elementos de la red UTRAN en este capítulo, y una descripción más extensa de las interfaces de la red en capítulos posteriores, se detallan las características principales de la red que también han sido requisitos para el diseño de la arquitectura, funciones y protocolos de la misma. Se pueden resumir en los siguientes puntos:

- ⇒ Soporte de UTRA y todas las funcionalidades relacionadas. En particular el mayor impacto del diseño de la red UTRAN ha sido el requerimiento de soportar el soft handover (un terminal

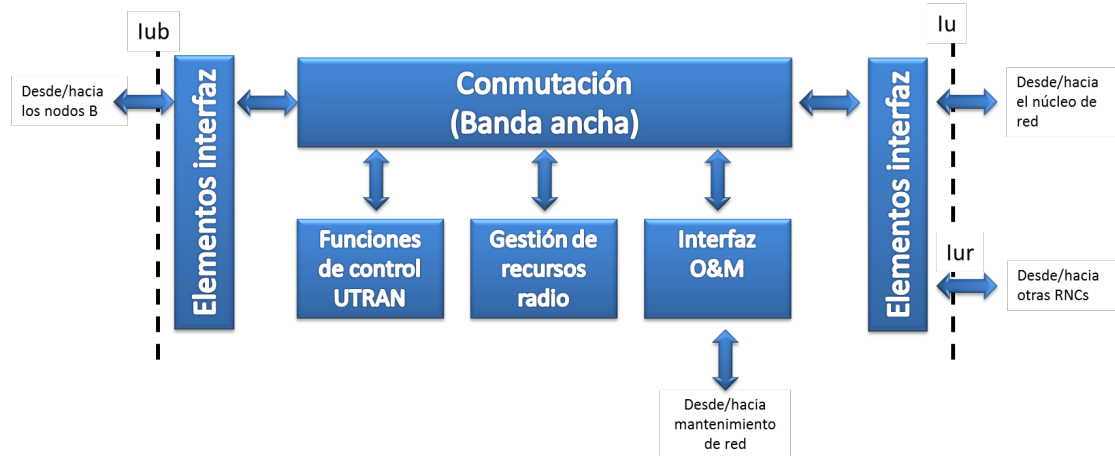


Figura 3.7: Arquitectura lógica básica de la RNC

conectado a la red a través de 2 o más celdas activas) y los algoritmos de gestión de recursos radio específicos de WCDMA.

- ⇒ Maximizar las similitudes en el tratamiento de datos de PS y de CS, con una única pila de protocolos estándar para la interfaz aire y con el uso de la misma interfaz para la conexión de la red UTRAN a los dominios de PS y CS del núcleo de red.
- ⇒ Maximizar las similitudes con GSM cuando sea posible.
- ⇒ Uso de ATM como principal tecnología de transporte en las redes UTRAN.
- ⇒ Uso del protocolo IP como mecanismo alternativo a partir de la Release 5.

3.3.1. La controladora RNC

La RNC es el elemento de la red responsable del control de los recursos radio en la red UTRAN. Sirve de unión entre el núcleo de red (normalmente a un MSC y un SGSN) y también termina el protocolo RRC (Radio Resource Control) que define los mensajes y procedimientos entre el teléfono y la red UTRAN. Se corresponde lógicamente con el BSC de las redes GSM.

La implementación de la RNC depende mucho del fabricante, aunque hay algunos puntos en común entre todos ellos. En la figura 3.7 se ilustra la arquitectura lógica básica de la RNC.

3.3.1.1. Papel lógico de la RNC

Dependiendo del escenario que se considere, la RNC puede desempeñar diferentes papeles.

Si la RNC controla un nodo B, se define como CRNC (*Controlling RNC*). La CRNC es responsable del control de carga y congestión de sus propias celdas, y también ejecuta el control de admisión y la provisión de códigos para los nuevos enlaces radio a establecer en dichas celdas.

En el caso de una conexión móvil-UTRAN en el que se utilicen recursos de más de un RNS (figura 3.8), las RNCs involucradas tienen dos roles lógicos diferenciados (con respecto a esta conexión móvil-UTRAN):

- ⇒ SRNC (*Serving RNC*). La SRNC para un terminal es la RNC que termina el enlace Iu para el transporte de datos de usuario y la correspondiente señalización RANAP (RAN Application

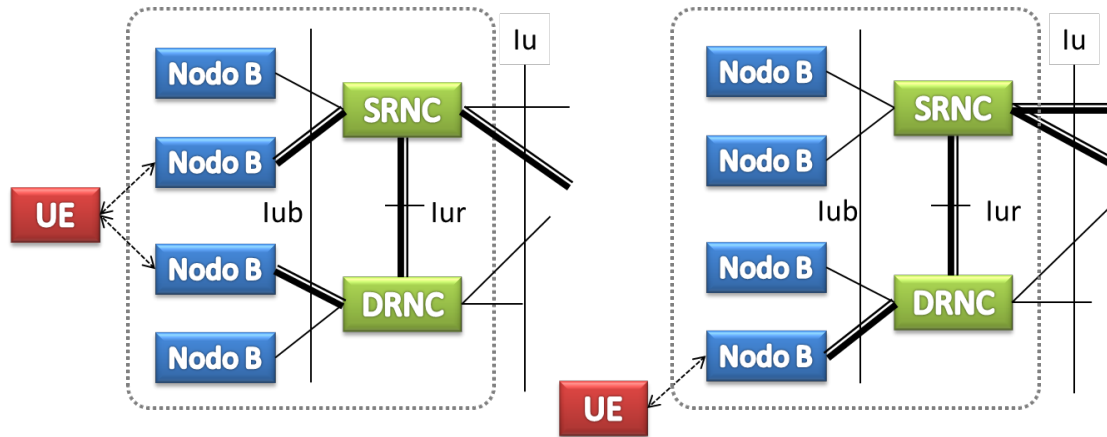


Figura 3.8: Roles lógicos de las RNCs para una conexión UE-UTRAN

Part) hacia o desde el núcleo de red. La SRNC también termina la señalización de RRC1, protocolo entre el UE y la UTRAN. Implementa el procesamiento L2 de los datos desde o hacia la interfaz radio. También se ejecutan las operaciones básicas de gestión de recursos radio, como el mapeo del RAB en la interfaz aérea, la decisión de handover y el control de potencia. Un UE conectado a la UTRAN tiene uno y sólo un SRNC.

⇒ DRNC (*Drift RNC*). La DRNC es cualquier RNC que controla las celdas que utiliza el terminal. El DRNC no desempeña el procesamiento L2 de los datos de usuario en plano, pero enruta los datos de forma transparente entre las interfaces Iub e Iur, excepto cuando el UE está usando un canal de transporte común o compartido. Un UE puede tener cero, uno o más DRNCs. Destacar que una RNC física normalmente contiene las funcionalidades de CRNC, SRNC y DRNC.

En la figura 3.8, el escenario de la izquierda muestra un UE en soft handover inter-RNC (la combinación se desarrolla en la SRNC). Por otro lado, el escenario de la derecha representa un UE que utiliza recursos sólo de un nodo B, controlado por la DRNC.

3.3.2. El Nodo B (estación base)

La función principal del Nodo B es el procesamiento de la interfaz aire en L1 (codificación y entrelazado del canal, adaptación de tasa, ensanchado, etc). También desempeña algunas funciones básicas de la gestión de recursos radio como el control de potencia. Se corresponde lógicamente con la estación base de las redes de GSM. El enigmático término de «nodo B» fue inicialmente adoptado como un término provisional durante el proceso de estandarización, pero nunca fue cambiado.

3.3.2.1. Estructura interna

El nodo B puede considerarse como la frontera de la red UTRAN y, en consecuencia, su tarea subyacente es la recepción y transmisión de las señales de radio, el filtrado de la señal y su amplificación, la modulación/demodulación de la señal e interactuar con el RAN.

La estructura interna de los nodos B depende mucho del fabricante, pero básicamente consiste en los componentes que se muestran en la figura 3.9, implementándose sus funciones por medio de

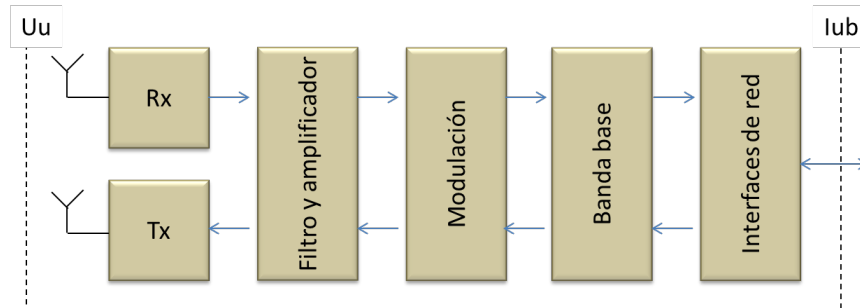


Figura 3.9: Estructura básica de un Nodo B

tarjetas. Su estructura a nivel lógico (es decir, cómo se integra el nodo dentro de la red UTRAN) es genérico.

Desde el punto de vista de la red radio y su control, el nodo B consiste en una serie de entidades lógicas llamadas «celdas». Una celda es la entidad más pequeña que puede considerarse en una red radio y tiene su número de identificación (Cell ID) que es visible por el terminal de usuario. Decir que se configura la red radio significa que son los datos de las celdas lo que en realidad se está cambiando.

Cada celda tiene un código de scrambling, y el terminal de usuario reconoce una celda por dos parámetros: el código de scrambling (cuando entra en el área de una celda) y el Cell ID (para la topología de la red). Una celda puede tener varios transmisores-receptores (TRXs). Un TRX mantiene los canales físicos a lo largo de la interfaz Uu y estos llevan los canales de transporte conteniendo la información que puede ser común o dedicada. Una celda contiene, como mínimo, un TRX. El TRX es una parte física del nodo B, que implementa varias funciones, como la conversión de los flujos de datos de la conexión terrestre Iub al canal radio y viceversa.

3.3.2.2. Técnica de modulación

Desde el punto de vista de sistema, la modulación que se emplee es un factor crítico ya que tiene una relación muy estrecha con la capacidad total del sistema y su rendimiento. WCDMA utiliza como técnicas de modulación, QPSK, su variante dual QPSK y QAM (16 QAM).

QAM, junto con las técnicas de modulación y codificación adaptativa (AMC) y HARQ son algunos de los elementos que se han añadido recientemente, junto con las mejoras de HSDPA.

La eficiencia de la modulación QAM, junto con la técnica de modulación/codificación adaptativa (AMC) incrementa el bit rate en el downlink en varios órdenes de magnitud. El nivel de modulación de QAM puede llegar hasta 64-QAM, pero debido a costes, complejidad y eficacia, se considera que una modulación de 16-QAM es más que razonable para una aplicación de HSDPA.

La figura 3.10 muestra las técnicas de modulación utilizadas. WCDMA utiliza las dos variantes de QPSK, la convencional para el enlace descendente y la dual en el ascendente, así como QAM en el downlink para aplicaciones de elevado bit-rate.

3.3.2.3. Funciones de control en el nodo B

Aunque las funciones principales del nodo B están relacionadas con la transmisión y recepción de la señal radio, así como otras funciones de procesamiento en banda base, hay no obstante otras funciones de control de la red UTRAN en las cuales el nodo se ve involucrado. Estas funciones

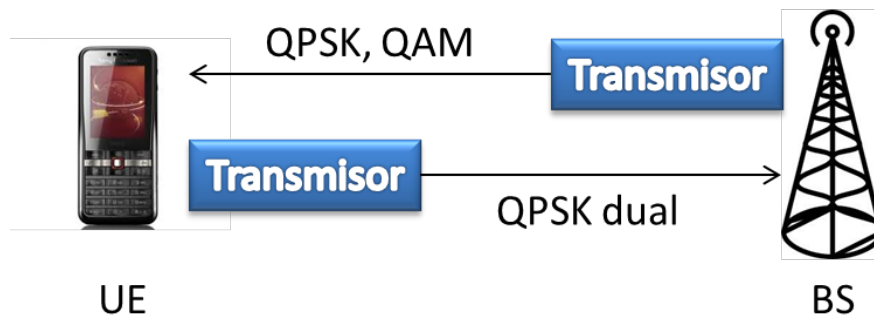


Figura 3.10: Técnicas de modulación utilizadas en WCDMA

están parcialmente limitadas para soportar otras funciones como recoger y filtrar medidas de radio y proporcionárselas a la RNC para ejecutar sus funciones de control.

Por otro lado, hay funciones en las que el nodo B es el principal responsable. Esto incluye la generación de códigos, ejecución de control de potencia y O&M (operaciones y mantenimiento). La función de O&M en el nodo incluye funciones a nivel «elemento de red» (Hardware y software) y a nivel de red UTRAN funciones lógicas controladas por la RNC. La función principal a nivel de «elemento de red», la cual es dependiente de la implementación, forma una base para el NMS (Network Management System). La manera en la que se lleva a cabo esta función depende del fabricante. Desde el punto de vista de la gestión de recursos radio (RRM), el nodo B está involucrado en el control de potencia (inner loop) y la generación de códigos, pero debido a las mejoras añadidas en la Release 5, también está involucrado en la entrega de paquetes. Según ha ido evolucionando al red UTRAN se han ido incrementando las funciones de control del nodo B. No obstante, la RNC todavía mantiene el papel central en términos de control absoluto de la red UTRAN.

3.3.3. Gestión de los recursos radio

Como se ha comentado anteriormente la gestión de los recursos radio (RRM) es responsabilidad de la red UTRAN. La gestión se localiza en el terminal de usuario, el nodo B y la RNC dentro de la red, tal y como se muestra en la figura 3.11. Hace uso de varios algoritmos, cuya meta es establecer la comunicación radio activando los criterios de QoS. Como protocolo para este fin se utiliza el RRC (Radio Resource Control).

Los algoritmos de gestión son los siguientes:

- ⇒ Control de handover.
- ⇒ Control de potencia.
- ⇒ Control de admisión (AC).
- ⇒ Planificador de paquetes (PS - *Packet Scheduling*).
- ⇒ Gestión de códigos.

3.4. Arquitectura de la red UTRAN de Ericsson

Ericsson dispone de diferentes equipos capaces de implementar el esquema de la red UMTS definido por la 3GPP y mostrado en la figura 3.12. En este apartado se describen los distintos

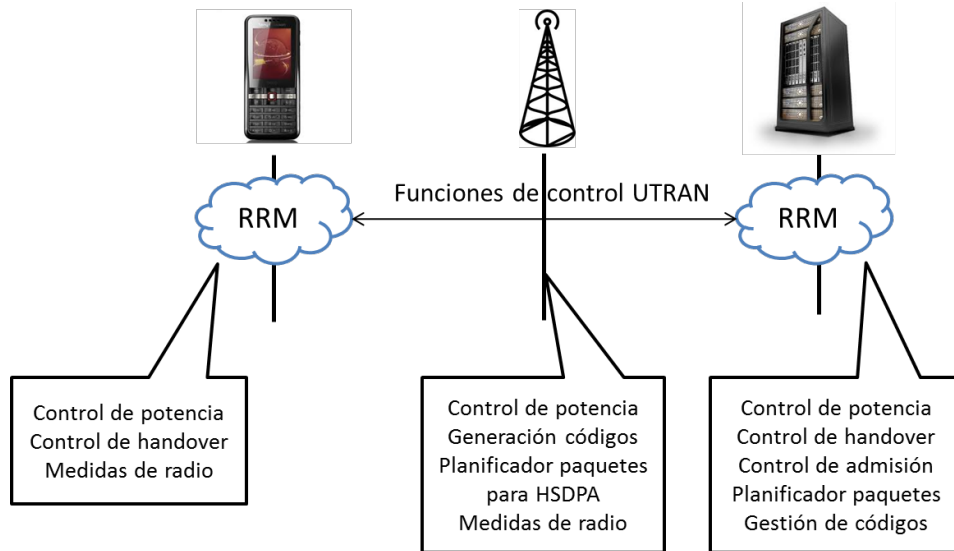


Figura 3.11: Principales entidades en la gestión de recursos radio

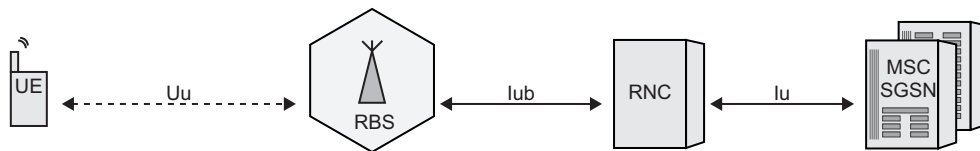


Figura 3.12: Esquema básico de la arquitectura UTRAN

elementos, dejando para el capítulo posterior las interfaces y la transmisión.

3.4.1. RNC

En la red a estudiar se dispone de una RNC 3810. Ésta se puede asemejar a un ordenador de alto rendimiento con las interfaces necesarias para poder conectarse con los nodos B y los puntos de conmutación MSC y SGSN. La arquitectura hardware del equipo está constituida por tarjetas de propósito general, donde se carga el control y los algoritmos, y tarjetas de propósito específico para manejar el tráfico. Están enlazadas por un conmutador ATM que también se conectan a unas tarjetas entrada/salida y con otras llamadas de tiempo para obtener la señal de reloj.

En términos de la capacidad de procesamiento y flujo de información que puede soportar, la RNC 3810 presenta las características mostradas en la tabla 3.1.

3.4.2. Nodo B

Ericsson denomina al nodo B como RBS (*Radio Bearer System*). En la red bajo estudio los nodos B se corresponden con el modelo RBS 3418. Estos equipos se componen de dos partes, por una lado, la *main unit* que corresponde al procesado en banda base y, por otro, a la *RRU* o cabeza remota que contempla la parte de radiofrecuencia y suele ir ubicada en el exterior. Así, permite que toda la potencia llegue a la antena y que el calor se disipe en el exterior, disminuyendo los requisitos de equipos de aire acondicionado y permitiendo la instalación con un menor espacio.

Ambos equipos se conectan por medio de fibra óptica, por lo que las distancias pueden ser grandes, con el único requisito de que la cabeza remota además de la fibra le llegue un cable de

Parámetro	Máximo
Mbps en Iub	175
Llamadas de voz	6400
Nº de nodos B	448
Celdas	1344
Relaciones de celda	127 (63 WCDMA+64GSM)
Vecindades externas	2304
Celdas externas GSM	7680
Áreas de localización	47
Áreas de enrutado	154
Áreas de servicio	2304
Contadores por servicio	1.700.000

Cuadro 3.1: Anchos de banda medio en el canal Iub

alimentación.

La unidad *main* puede estar alojada tanto en racks de interior como de intemperie, en cualquier caso, ocupa 4 Us de un rack. En estas 4 Us encontramos 11 slots donde se pueden alojar diferentes tarjetas para procesar la señal en banda base y gobernar el funcionamiento del nodo B.

El esquema final de la configuración del nodo B se muestra en la figura 3.13.

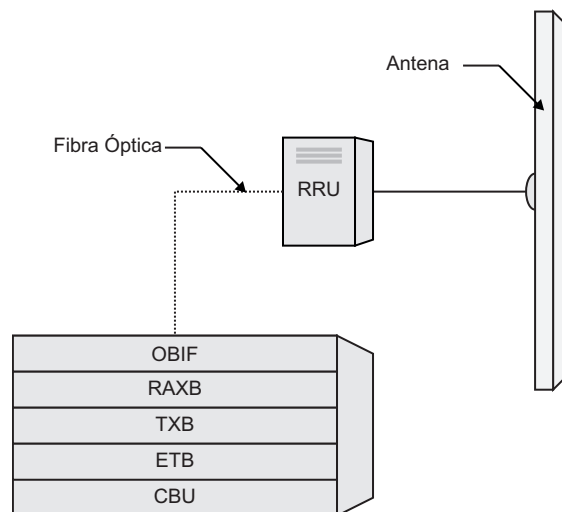


Figura 3.13: Esquema de componentes de un Nodo B Ericsson (RBS)

Las tarjetas que componen la *main unit* son cinco:

⇒ *Control Base Unit* (CBU), es la unidad de control central de nodo B. Ejecuta la mayor parte de las funciones de la RBS y controla las tarjetas que se han instalado en la *main unit*. La CBU contiene, además, un conmutador ATM que se comunica con las otras tarjetas a través del *backpanel* y proporciona una interfaz IMA de hasta 4 E1s. Dispone también de un puerto para dotar al nodo de una referencia de reloj externa a través de la propia transmisión o mediante GPS, que estabiliza y redistribuye el resto de tarjetas. Por último, se encarga de estabilizarl y alimentar a estas últimas.

⇒ ETB. Proporciona un extra de puertos para comunicación externa. Son tarjetas opcionales. Pueden ser de diferentes tipos, ETMC o ETMFX, según la interfaz que implemente, ATM o IP.

- ⇒ RAXB. Se encarga de la parte receptora del nodo. Las principales tareas son, demodulación de la señal mediante receptor RAKE, estimación de canal y combinación de máxima similitud, detección de acceso aleatorio, decodificación y desentrelazado. Se pueden instalar varias tarjetas RAXB. Se consigue cierta redundancia por la compartición de carga de las mismas. Existen principalmente dos tipos, RAXB12 y RAXB13, que se diferencian en la capacidad que ofrecen, 64 y 128 channel elements respectivamente.
- ⇒ TXB. Se encarga de la parte transmisora del nodo. Se encargan de la codificación, modulación y expansión del espectro y procesado en banda base. Se pueden instalar hasta dos tarjetas. Las que son capaces de gestionar HSPA se denominan HS-TX y existen dos tipos: HS-TX45 y HS-TX60; se diferencian en la capacidad de códigos máxima que pueden tratar, 45 y 60 respectivamente.
- ⇒ OBIF. Esta tarjeta proporciona la conexión óptica con las RRUs. Se permite la conexión con hasta 6 RRUs.

De acuerdo al tipo de tarjeta que se instale en el nodo B se dispondrá de más o menos capacidad de procesamiento (CE y códigos) y se podrán gestionar más o menos conexiones. Será, pues, un «cuello de botella» a tener en cuenta, al igual que lo son los códigos de ensanchamiento o *spreading factors* (SF).

3.5. Tecnología LTE (*UMTS Long Term Evolution*)

3.5.1. Introducción

En este apartado se describe la arquitectura de una red de comunicaciones móviles basada en las especificaciones del sistema LTE. Para ello, a partir del 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 termino 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.

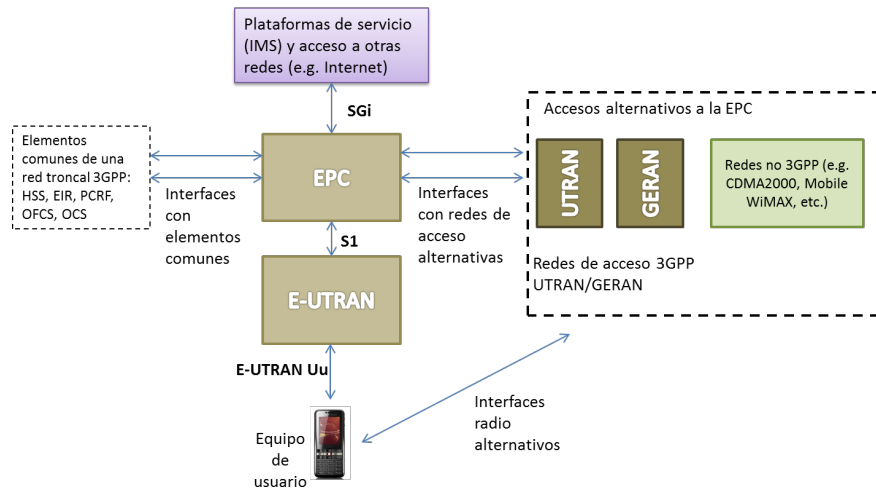


Figura 3.14: Arquitectura del sistema LTE

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 esta Tesis se ha adoptado también este criterio.

3.5.2. Arquitectura del sistema LTE

Atendiendo a la arquitectura general de los sistemas 3GPP, en la Figura 3.14 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 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 (ERAB).

En la Figura 3.14 se muestran las principales interfaces de E-UTRAN y EPC. 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 Uu. 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 S-Gi de la EPC. La interfaz S-Gi 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 3.14 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 suscripción de sus usuarios (HSS).

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.

Finalmente, aunque no quede reflejado en la Figura 3.14, 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.

3.5.3. Arquitectura de E-UTRAN

La arquitectura de la red de acceso se compone de una única entidad de red 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

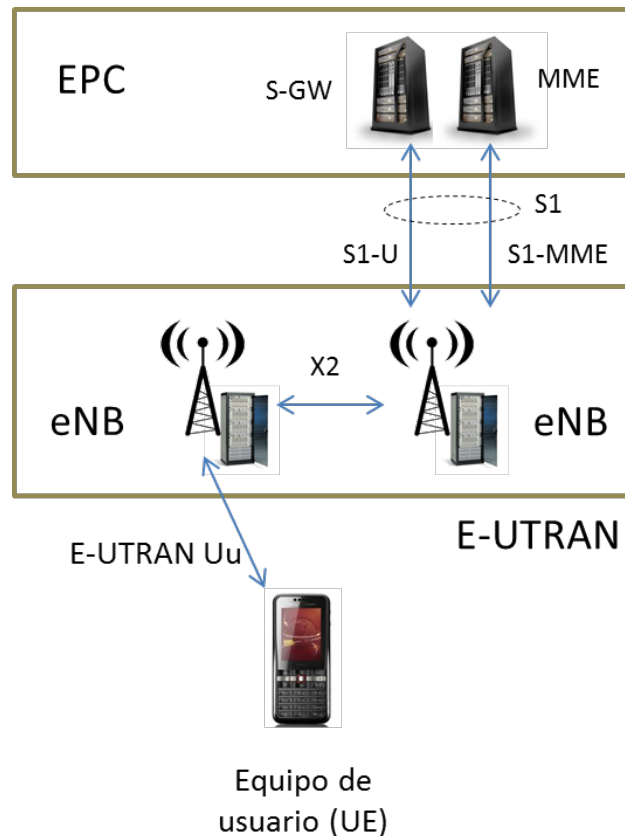


Figura 3.15: Red de acceso E-UTRAN

y TS 36.401.

Tal y como se ilustra en la Figura 3.15, 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.

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. 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*.

3.5.4. Entidades de red: *Evolved Node B* (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.

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-fl ex).

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

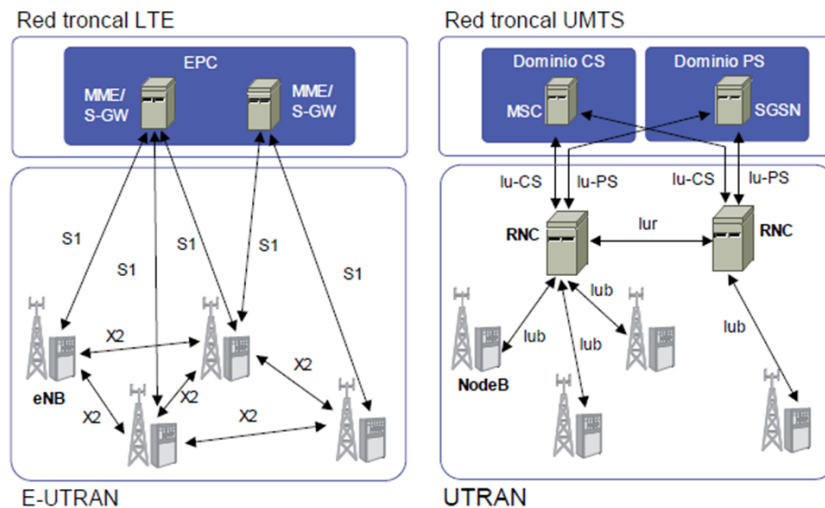


Figura 3.16: Comparativa de la arquitectura de red de E-UTRAN y UTRAN

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.

3.6. 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 3.16 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 con objeto de comparar ambas arquitecturas.

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 25.40. 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 Iub 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 sí mediante interfaces específicas como la interfaz Iur 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 Iu-PS, entre RNCs y los nodos SGSNs del dominio de paquetes, y Iu-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 3.16 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 procesamiento 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. 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 procesamiento 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).

3.7. Descripción de las interfaces

Las redes UMTS se componen de dos entidades que pueden tratarse separadamente: por un lado, la red de acceso y, por otro, el núcleo de red. En este caso se analiza la parte correspondiente a la red de acceso y, más concretamente, una red UTRAN del fabricante sueco Ericsson. Se basa en la configuración correspondiente a la versión P6 de este fabricante, la cual implementa el estándar UMTS Release 6 de la 3GPP. En este apartado se pretende resumir brevemente las distintas interfaces que se definen en UMTS, explicando qué elementos de la red enlazan y qué función tiene cada una de ellas. Dentro de la red UTRAN se definen dos tipos de interfaces (internos a UTRAN), la interfaz «Iub» y la interfaz «Iur» (opcional). Por otra parte, la UTRAN se comunica con los equipos de usuario (UE) y los nodos de acceso al núcleo de red (MSC o SGSN según el dominio de que se trate) a través de las interfaces Uu e Iu, respectivamente. En el Apéndice E se detalla la arquitectura de protocolos de la red UTRAN, la arquitectura por capas de las interfaces Iu, Iur e Iub, y se concretan los protocolos de cada una de ellas. Por último se hace mención a los protocolos de radio, entre los cuales cabe destacar el protocolo RRC (*Radio Resource Control*).

3.7.1. Interfaz Iub (Iu bis)

Esta interfaz conecta a cada nodo B con la RNC que le controla (CRNC). Esta interfaz puede considerarse como una prolongación de los canales de transporte de la interfaz radio. Dicho de otra manera, la interfaz Iub es responsable del transporte de las tramas MAC entre el Nodo B y la RNC.

3.7.2. Interfaz Iur

Permite la comunicación entre dos RNC para el soporte de trasposos de soft-handover entre nodos B pertenecientes a distintos subsistemas RNS. Durante este tipo de trasposos, la interfaz Iur permite el transporte de las tramas MAC entre la RNC auxiliar (DRNC) y la RNC servidor (SRNC). La interfaz Iur no tiene equivalente en el subsistema radio de GSM.

3.7.3. Interfaz Iu

La interfaz Iu es el nexo de unión entre la red de acceso radio y el núcleo de red, constituyendo el tercer eslabón (después de la interfaz radio y la interfaz Iub) en la cadena que une a los equipos de usuario con la red UMTS. Como se muestra en la figura 3.6, la interfaz Iu se descompone en dos interfaces lógicas, «Iu-CS» e «Iu-PS», que ligan la red de acceso radio a los dos dominios del núcleo de red (CS y PS, respectivamente). En otras palabras, Iu-CS es la interfaz que une al RNC con la MSC de la que depende, mientras que Iu-PS lo liga al SGSN. Las interfaces Iu-CS e Iu-PS son equivalentes a las interfaces A y Gb de las redes GSM/GPRS.

3.7.4. Interfaz Uu

Finalmente, la UTRAN se comunica con los equipos de usuario mediante la interfaz radio Uu. Como se sabe, se trata de una interfaz basada en W-CDMA con dos modos de operación (FDD y TDD), sobre el que se definen los distintos canales radio.

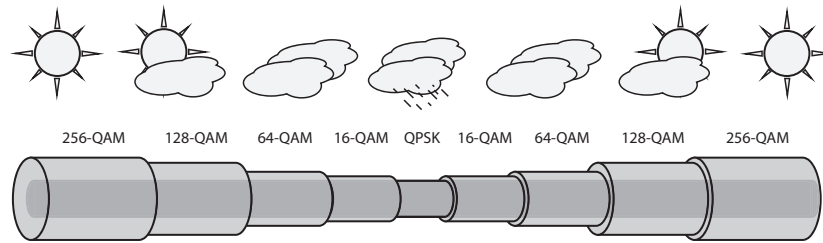


Figura 3.18: Funcionamiento de la modulación adaptativa

de tramas IP (capa 3).

Radioenlaces PDH/SDH

En la red UTRAN bajo estudio se encuentran instalados radioenlaces PDH y SDH de los fabricantes Ericsson y Siemens. En concreto, enlaces PDH Mini-Link E de Ericsson en configuraciones de 4x2, 8x2, 16x2 y 32x2 E1s y enlaces SraL y SraL XD de Siemens en configuraciones similares. Para enlaces SDH, la red tiene enlaces Traffic Node de Ericsson con capacidades de 1xSTM-1 o 2xSTM-1.

Radioenlaces Ethernet

En cuanto a los radioenlaces Ethernet, la red tiene instalados enlaces de los fabricantes Nokia-Siemens y Huawei, en concreto, de los modelos FlexiPacket y RTN respectivamente. Cabe destacar, que este tipo de radioenlaces tienen un importante impacto a la hora de dimensionar el canal Iub, no solo porque sean partes del mismo, sino por su propio funcionamiento, ya que ambos modelos incorporan «modulación adaptativa». Esta tecnología permite adaptar la modulación empleada durante la transmisión/recepción en virtud de la calidad radio que se tengan medida a partir de la tasa de bit erróneos (BER). Por tanto, este tipo de enlaces harán variar la capacidad del Iub en virtud de las condiciones radio (que vendrán en la mayoría de los casos marcadas por las condiciones climáticas) haciendo que en determinadas condiciones la capacidad disponible no sea fija, como sucede en otros tipos de transmisión (fibra óptica, cables de pares, enlaces PDH/SDH), tal y como se ilustra en la figura 3.18.

	4 QAM	16 QAM	32 QAM	64 QAM	128 QAM	256 QAM
7 MHz	9	17	N/A	27	32	38
14 MHz	17	35	N/A	57	68	80
28 MHz	43	86	N/A	133	155	179
56 MHz	85	169	N/A	260	305	351

Cuadro 3.2: Capacidad radioenlaces Nokia-Siemens FlexiPacket vs. ancho de banda del canal y modulación empleados

3.8.2. xDSL

Una buena alternativa en ciudades es dotar a aquellas estaciones próximas a centrales de telefonía fija transmisión mediante xDSL a través de cable de pares. Este tipo de transmisión permite dotar a los nodos B de conectividad IP a tasas de que oscilan entre los 20 y los 30 Mbps, valores inferiores a los que se pueden obtener mediante radioenlaces y muy por debajo de la fibra óptica.

	4 QAM	16 QAM	32 QAM	64 QAM	128 QAM	256 QAM
7 MHz	10	20	25	32	38	44
14 MHz	20	42	51	66	78	90
28 MHz	42	84	105	133	158	183
56 MHz	84	168	208	265	313	363

Cuadro 3.3: Capacidad radioenlaces Huawei PTN vs. ancho de banda del canal y modulación empleados

3.8.3. Fibra Óptica

La red UTRAN de Ericsson bajo estudio tiene conexión mediante fibra óptica en nodos B que aglutinan grandes cantidades de tráfico. La transmisión mediante fibra óptica dota a los nodos de una capacidad máxima de 1 Gbps. Sin embargo, el valor real máximo es inferior, debido a los múltiples cuellos de botella que existen, como son el cableado utilizado, la configuración de los puertos, los equipos con tasas de *switching* inferiores a 1 Gbps o la propia interfaz radio entre otros. Es por ello por lo que en la red en estudio la fibra desplegada es multimodo trabajando en segunda ventana (1000-1300 nm).

3.8.4. Backhaul

Superado el tramo de última milla de la red, existe la red de transporte que irá agregando tráficos de última milla hasta enlaces cada vez más troncales que acaban llegando a la RNC. Para realizar estas agregaciones son necesarios una series de equipos: SMX y PTNs. En la red bajo estudio los SMX se corresponden con equipos Surpass hIT familia 7000 de Nokia-Siemens Networks, mientras que los PTNs son equipos OptiX PTN 3900 del fabricante chino Huawei.

Los equipos SMX permiten realizar multiplexaciones de flujos PDH/SDH e incluso Ethernet; es decir, pueden, por ejemplo, agregar tráficos procedentes de diversos contenedores PDH y obtener una única trama SDH de salida.

Los equipos PTN son una de las últimas tecnologías introducidas en las redes de transmisión. En realidad, son nodos MPLS; gracias a ellos se pueden unificar el servicio de transporte de datos para las redes basadas en circuitos y las basadas en paquetes. Puede ser utilizado para transportar diferentes tipos de tráfico, incluyendo tráfico de voz y de paquetes IP.

3.9. Transmisión, capa de red

3.9.1. ATM

ATM (*Asynchronous Transfer Mode*) es una tecnología de transporte ampliamente empleada en redes de telecomunicaciones y que opera a nivel de capa 2 según el modelo de referencia OSI.

ATM es una tecnología orientada a conexión, diseñada para el transporte de múltiples tipos de servicios entre ellos voz, vídeo y datos a través de redes públicas y privadas y provee gran escalabilidad del ancho de banda.

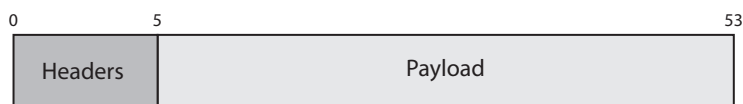


Figura 3.19: Celda ATM

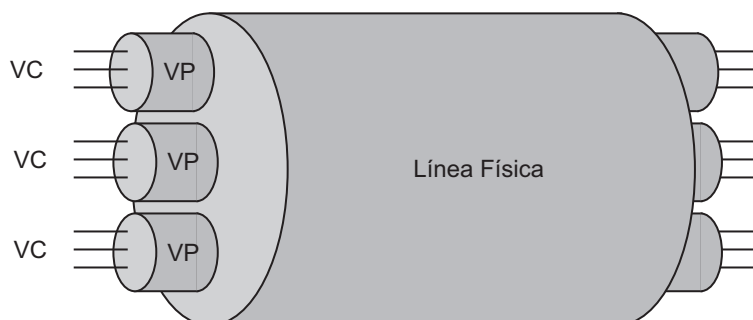


Figura 3.20: Conexiones virtuales en ATM

3.9.1.1. Conceptos

Celdas

ATM se basa en la multiplexación de celdas. Son pequeños paquetes de tamaño fijo, estandarizadas a 53 bytes, figura B.1. De estos 53 bytes los 5 primeros se corresponden con datos de cabecera y el resto con la carga de datos a transportar. El hecho de que las celdas sean de tamaño fijo facilita el proceso de conmutación al no tener que verificar el tamaño variable de celdas. No hay que incluir en el encabezado el tamaño de la celda. El porqué de los 53 bytes se explica debido a que celdas más grandes provocan tiempos de retardo en la transmisión que servicios como voz y vídeo no pueden tolerar, mientras que celdas más pequeñas reducen la eficiencia en el manejo del ancho de banda.

ATM toma los paquetes o tramas de los servicios a transportar y los divide en pequeños segmentos que multiplexa a través de una línea.

El encargado de segmentar los datos es la capa de adaptación ATM o AAL.

Conexiones Virtuales

Las conexiones virtuales son conexiones lógicas para establecer comunicación entre dos clientes de la red ATM. Varias conexiones virtuales subdividen una sola interfaz física.

El concepto de canal virtual representa una conexión entre dos puntos de la red ATM mientras que la trayectoria virtual o *virtual path* representa un conjunto de canales virtuales (ver figura B.2).

Conmutación en ATM

A la hora de hablar de tipos de conmutación pueden encontrarse dos tipos:

Conmutación de circuitos

Dedica un circuito físico a cada comunicación. Su mayor ventaja es poseer tiempos de retraso pequeños y estables en la transmisión, además de ofrecer un ancho de banda garantizado a los

usuarios. En contrapartida las conexiones son poco flexibles y desaprovecha gran parte del ancho de banda de la línea utilizada.

Conmutación de paquetes

Se comparte el medio físico, tal y como sucede por ejemplo en la tecnología IP. Su mayor ventaja es que aprovecha al máximo el ancho de banda y dota de mayor flexibilidad a las conexiones. En contrapartida no se garantiza el ancho de banda y el retraso de los paquetes es variable.

ATM combina las ventajas de ambos métodos de conmutación:

- ⇒ Ancho de banda garantizado y el tiempo de retraso estable de la conmutación de circuitos.
- ⇒ Aprovechamiento del ancho de banda y la flexibilidad en las conexiones.

3.9.1.2. Calidad de Servicio

ATM implementa un fuerte control sobre las conexiones a fin de dotar de calidad de servicio a cada una. Para ello, cuando se establece una nueva conexión se inicia un «contrato» entre la red y el usuario por el que ambos se comprometen a un intercambio de información dado. La red reserva los recursos necesarios y el usuario se compromete a transmitir de acuerdo a unos criterios:

- ⇒ Se especifican parámetros para el flujo de datos de una conexión.
- ⇒ Establece un compromiso de la red hacia las terminales.

La calidad de servicio vendrá dada por el tipo de tráfico que se curse. Existen diferentes categorías:

⇒ Flujo constante CBR:

- Ancho de banda estable y tiempo de retraso garantizado.
- Transporta servicios que requieren señal de sincronía.
- Utilizado para voz y vídeo digitales y emulación de circuitos.

⇒ Flujo variable - Tiempo Real VBR-RT:

- Provee ancho de banda variable por demanda a servicios susceptibles al tiempo de transmisión.
- Utilizado para voz y vídeo «paquetizados».

⇒ Flujo variable VBR-NRT:

- Provee ancho de banda variable por demanda a servicios menos susceptibles al tiempo de transmisión.

⇒ Inespecificado UBR:

- Provee el mejor trato posible sin garantías de ninguna especie.

⇒ Flujo disponible ABR:

- Provee de un ancho de banda variable de acuerdo a la congestión de la red. \item Se utiliza para transportar servicios muy flexibles al tiempo de transmisión y al ancho de banda como IP y IPX.
- Requiere de Control de Flujo a lazo cerrado.

3.9.1.3. Capas de Adaptación

AAL nace como un mecanismo de la tecnología ATM para la adaptación del formato de celda desde la capa de enlace de datos a los niveles superiores de otras arquitecturas, tales como TCP/IP o el modelo OSI. Está definido en el ITU y se diseñaron 5 protocolos adaptados a las diferentes arquitecturas de red que pueden soportar ATM.

Constan de dos partes:

⇒ Subcapa de convergencia o CS (*Convergence Sub-layer*).

⇒ Subcapa de segmentación y reensamble o SAR (*Segmentation and Reassembly Sub-layer*).

Capa AAL1

Se emplea habitualmente con conexiones que transporten un flujo constante de bits, por lo que es idónea para el transporte de voz y vídeo. Transporta servicios que requiera sincronización entre emisor y receptor y además puede emular circuitos.

Capa AAL2

AAL Tipo 2 soporta rt-VBR (*Variable Bit Rate in real time*), de circuito orientado a la conexión de tráfico síncrono. Su servicio es de baja prioridad y garantizado. Se utiliza en compresión de vídeo.

Capa AAL3-4

Soportan al VBR, tráfico de datos, circuitos orientados a la conexión, tráfico asíncrono (por ejemplo X.25 de datos) o a los paquetes de datos no orientados a la conexión (ej: tráfico SMDS) con una cabecera (*header*) adicional de 4 bytes en el *payload* de la celda. Por ejemplo, Frame Relay y X.25. Su servicio es de alta prioridad y no garantizado.

Capa AAL5

Este AAL ha sido diseñado para utilizarse bajo TCP/IP y está normalizado en la RFC 1577. AAL Tipo 5 es similar a AAL 3/4 con un programa de simplificación de cabecera (*header*) de información. Este AAL asume que los datos son secuenciales desde el usuario final y usa el bit *Payload Type Indicator* (PTI) para indicar la última celda en transmitirse. Ejemplos de este servicio son el clásico IP sobre ATM, Ethernet sobre ATM, SMDS, y emulación LAN (LANE). Su servicio es de baja prioridad y no garantizado.

3.9.2. MPLS

Introducción

El enorme crecimiento de la red Internet ha convertido al protocolo IP (Internet Protocol) en la base de las actuales redes de telecomunicaciones, contando con más del 80 % del tráfico cursado. La versión actual de IP, conocida por IPv4 y recogida en la RFC 791, lleva operativa desde 1980. Este protocolo de capa de red (Nivel 3 OSI), define los mecanismos de la distribución o encaminamiento de paquetes, de una manera no fiable y sin conexión, en redes heterogéneas; es decir, únicamente está orientado a servicios no orientados a conexión y a la transferencia de datos, por lo que se

suele utilizar junto con TCP (Transmission Control Protocol) (Nivel 4 de OSI) para garantizar la entrega de los paquetes.

A mediados de la década de los 90, la demanda por parte de los clientes de los ISP (Internet Service Providers) de aplicaciones multimedia con altas necesidades de ancho de banda y una calidad de servicio o QoS (Quality of Service) garantizada, propiciaron la introducción de ATM (Asynchronous Transfer Mode) en la capa de enlace (Nivel 2 de OSI) de sus redes. En esos momentos, el modelo de IP sobre ATM satisfacía los requisitos de las nuevas aplicaciones, utilizando el encaminamiento inteligente de nivel 3 de los routers IP en la red de acceso, e incrementando el ancho de banda y rendimiento basándose en la alta velocidad de los conmutadores de nivel 2 y los circuitos permanentes virtuales de los switches ATM en la red troncal. Esta arquitectura, no obstante, presenta ciertas limitaciones, debido a: la dificultad de operar e integrar una red basándose en dos tecnologías muy distintas, la aparición de switches ATM e IP de alto rendimiento en las redes troncales, y la mayor capacidad de transmisión ofrecida por SDH/SONET (Synchronous Digital Hierarchy / Synchronous Optical Network) y DWDM (Dense Wavelength Division Multiplexing) respecto a ATM.

Durante 1996, empezaron a aparecer soluciones de conmutación de nivel 2 propietarias diseñadas para el núcleo de Internet que integraban la conmutación ATM con el encaminamiento IP; como por ejemplo, Tag Switching de Cisco o Aggregate Route-Based IP Switching de IBM. La base común de todas estas tecnologías, era tomar el software de control de un router IP, integrarlo con el rendimiento de reenvío con cambio de etiqueta de un switch ATM y crear un router extremadamente rápido y eficiente en cuanto a coste. La integración en esta arquitectura era mayor, porque se utilizaban protocolos IP propietarios para distribuir y asignar los identificadores de conexión de ATM como etiquetas; pero los protocolos no eran compatibles entre sí y requerían aún de infraestructura ATM.

Finalmente en 1997, el IETF (Internet Engineering Task Force) establece el grupo de trabajo MPLS (MultiProtocol Label Switching) para producir un estándar que unificase las soluciones propietarias de conmutación de nivel 2. El resultado fue la definición en 1998 del estándar conocido por MPLS, recogido en la RFC 3031. MPLS proporciona los beneficios de la ingeniería de tráfico del modelo de IP sobre ATM, pero además, otras ventajas; como una operación y diseño de red más sencillo y una mayor escalabilidad. Por otro lado, a diferencia de las soluciones de conmutación de nivel 2 propietarias, está diseñado para operar sobre cualquier tecnología en el nivel de enlace, no únicamente ATM, facilitando así la migración a las redes ópticas de próxima generación, basadas en infraestructuras SDH/SONET y DWDM.

Concepto de MPLS

MPLS es un estándar IP de conmutación de paquetes del IETF, que trata de proporcionar algunas de las características de las redes orientadas a conexión a las redes no orientadas a conexión. En el encaminamiento IP sin conexión tradicional, la dirección de destino junto a otros parámetros de la cabecera, es examinada cada vez que el paquete atraviesa un router. La ruta del paquete se adapta en función del estado de las tablas de encaminamiento de cada nodo, pero, como la ruta no puede predecirse, es difícil reservar recursos que garanticen la QoS; además, las búsquedas en tablas de encaminamiento hacen que cada nodo pierda cierto tiempo, que se incrementa en función de la longitud de la tabla.

Sin embargo, MPLS permite a cada nodo, ya sea un switch o un router, asignar una etiqueta a

cada uno de los elementos de la tabla y comunicarla a sus nodos vecinos. Esta etiqueta es un valor corto y de tamaño fijo transportado en la cabecera del paquete para identificar un FEC (Forward Equivalence Class), que es un conjunto de paquetes que son reenviados sobre el mismo camino a través de la red, incluso si sus destinos finales son diferentes. La etiqueta es un identificador de conexión que sólo tiene significado local y que establece una correspondencia entre el tráfico y un FEC específico. Dicha etiqueta se asigna al paquete basándose en su dirección de destino, los parámetros de tipo de servicio, la pertenencia a una VPN, o siguiendo otro criterio. Cuando MPLS está implementado como una solución IP pura o de nivel 3, que es la más habitual, la etiqueta es un segmento de información añadido al comienzo del paquete. Los campos de la cabecera MPLS de 4 bytes, son los siguientes:

Label (20 bits). Es el valor actual, con sentido únicamente local, de la etiqueta MPLS. Esta etiqueta es la que determinará el próximo salto del paquete. CoS (3 bits). Este campo afecta a los algoritmos de descarte de paquetes y de mantenimiento de colas en los nodos intermedios, es decir, indica la QoS del paquete. Mediante este campo es posible diferenciar distintos tipos de tráfico y mejorar el rendimiento de un tipo de tráfico respecto a otros. Stack (1 bit). Mediante este bit se soporta una pila de etiquetas jerárquicas, es decir, indica si existen más etiquetas MPLS. Las cabeceras MPLS se comportan como si estuvieran apiladas una sobre otra, de modo que el nodo MPLS tratará siempre la que esté más alto en la pila. La posibilidad de encapsular una cabecera MPLS en otras, tiene sentido, por ejemplo, cuando se tiene una red MPLS que tiene que atravesar otra red MPLS perteneciente a un ISP u organismo administrativo externo distinto; de modo que al terminar de atravesar esa red, se continúe trabajando con MPLS como si no existiera dicha red externa.

Elementos de una red MPLS

En MPLS un concepto muy importante es el de LSP (Label Switch Path), que es un camino de tráfico específico a través de la red MPLS, el cual se crea utilizando los LDPs (Label Distribution Protocols), tales como RSVP-TE (ReSerVation Protocol – Traffic Engineering) o CR-LDP (Constraint-based Routing – Label Distribution Protocol); siendo el primero el más común. El LDP posibilita a los nodos MPLS descubrirse y establecer comunicación entre sí con el propósito de informarse del valor y significado de las etiquetas que serán utilizadas en sus enlaces contiguos. Es decir, mediante el LDP se establecerá un camino a través de la red MPLS y se reservarán los recursos físicos necesarios para satisfacer los requerimientos del servicio previamente definidos para el camino de datos.

Una red MPLS está compuesta por dos tipos principales de nodos, los LER (Label Edge Routers) y los LSR (Label Switching Routers), tal y como se muestra en el ejemplo de la Figura 1. Los dos son físicamente el mismo dispositivo, un router o switch de red troncal que incorpora el software MPLS; siendo su administrador, el que lo configura para uno u otro modo de trabajo. Los nodos MPLS al igual que los routers IP normales, intercambian información sobre la topología de la red mediante los protocolos de encaminamiento estándar, tales como OSPF (Open Shortest Path First), RIP (Routing Information Protocol) y BGP (Border Gateway Protocol), a partir de los cuales construyen tablas de encaminamiento basándose principalmente en la alcanzabilidad a las redes IP destinatarias. Teniendo en cuenta dichas tablas de encaminamiento, que indican la dirección IP del siguiente nodo al que le será enviado el paquete para que pueda alcanzar su destino final, se establecerán las etiquetas MPLS y, por lo tanto, los LSP que seguirán los paque-

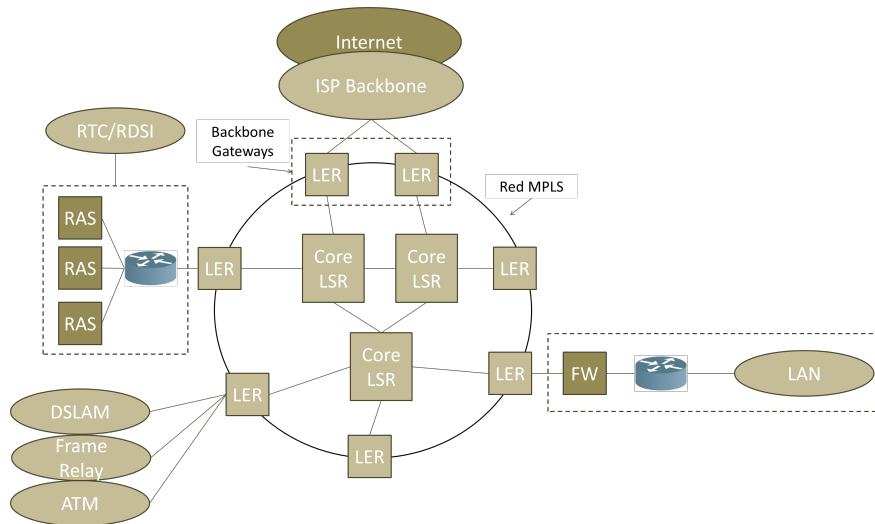


Figura 3.21: Ejemplo de una red MPLS

tes. No obstante, también pueden establecerse LSP que no se correspondan con el camino mínimo calculado por el protocolo de encaminamiento.

Los LERs están ubicados en el borde de la red MPLS para desempeñar las funciones tradicionales de encaminamiento y proporcionar conectividad a sus usuarios, generalmente routers IP convencionales. El LER analiza y clasifica el paquete IP entrante considerando hasta el nivel 3, es decir, considerando la dirección IP de destino y la QoS demandada; añadiendo la etiqueta MPLS que identifica en qué LSP está el paquete. Es decir, el LER en vez de decidir el siguiente salto, como haría un router IP normal, decide el camino entero a lo largo de la red que el paquete debe seguir. Una vez asignada la cabecera MPLS, el LER enviará el paquete a un LSR. Los LSR están ubicados en el núcleo de la red MPLS para efectuar encaminamiento de alto rendimiento basado en la conmutación por etiqueta, considerando únicamente hasta el nivel 2. Cuando le llega un paquete a una interfaz del LSR, éste lee el valor de la etiqueta de entrada de la cabecera MPLS, busca en la tabla de conmutación la etiqueta e interfaz de salida, y reenvía el paquete por el camino predefinido escribiendo la nueva cabecera MPLS. Si un LSR detecta que debe enviar un paquete a un LER, extrae la cabecera MPLS; como el último LER no conmuta el paquete, se reducen así cabeceras innecesarias.

Implementaciones de MPLS

Una vez visto el concepto de MPLS, veamos los distintos tipos de implementaciones actuales, en concreto: MPLS como una solución IP sobre Ethernet, IP sobre ATM, e IP sobre Frame Relay. No se contempla la aplicación de MPLS a las redes ópticas de próxima generación, conocida como GMPLS (Generalized MPLS), por encontrarse aún en proceso de estudio y estandarización por parte del IETF. GMPLS es una extensión natural de MPLS para ampliar el uso de MPLS como un mecanismo de control y provisión, no únicamente de caminos en dispositivos basados en paquetes, sino también de caminos en dispositivos no basados en paquetes; como los conmutadores ópticos de señales multiplexadas por división en longitud de onda, los conmutadores de fibras ópticas, y los conmutadores de señales digitales multiplexadas por división en el tiempo. Es decir, GMPLS busca una integración total en la parte de control de las redes de conmutación de paquetes IP y las redes ópticas SONET/SDH y DWDM; dando lugar a las redes ópticas inteligentes de próxima

generación, cuya evolución final será la integración de IP directamente sobre DWDM utilizando algún mecanismo de encapsulamiento como los “digital wrappers”.

La implementación de MPLS como una solución IP sobre Ethernet, Fast Ethernet o Gigabit Ethernet, es la conocida como IP pura. Puesto que IPv4 es un protocolo diseñado mucho antes que MPLS, en este caso, la etiqueta MPLS está ubicada después de la cabecera de nivel 2 y antes de la cabecera IP. Los LSR saben como conmutar utilizando la etiqueta MPLS en vez de utilizar la cabecera IP. El funcionamiento de IPv4 ha sido totalmente satisfactorio, no obstante, el sorprendente crecimiento de Internet evidenció importantes carencias, como: la escasez de direcciones IP, la imposibilidad de transmitir aplicaciones en tiempo real y los escasos mecanismos de seguridad. Estas limitaciones propiciaron el desarrollo de la siguiente generación del protocolo Internet o IPv6, definido en la RFC 1883. La versión IPv6 puede ser instalada como una actualización del software en los dispositivos de red de Internet e interoperar con la versión actual IPv4, produciéndose esta migración progresivamente durante los próximos años. En este caso, la etiqueta MPLS forma parte de la propia cabecera IPv6, estando su uso descrito en la RFC 1809.

La implementación de MPLS como una solución IP sobre ATM también está muy extendida. Primeramente indicar, que MPLS no fue desarrollado para reemplazar ATM, sino para complementarlo. De hecho, la aparición de switches ATM e IP con soporte de MPLS, ha integrado las ventajas de los routers IP y los switches ATM y ha supuesto una mejora de la relación precio/rendimiento de estos dispositivos. La diferencia principal entre MPLS y otras soluciones de IP sobre ATM, es que las conexiones MPLS se establecen utilizando LDP, y no por los protocolos de señalización ATM tradicionales, tales como PNNI (Private Network to Network Interface). Por otro lado, MPLS elimina la complejidad de hacer corresponder el direccionamiento IP y la información de encaminamiento directamente en las tablas de conmutación de ATM, puesto que LDP entiende y utiliza direcciones IP y los protocolos de encaminamiento utilizados en las redes MPLS son los mismos que los utilizados en las redes IP. En este caso, descrito en la RFC 3035, la etiqueta es el valor del VPI/VCI (Virtual Path Identifier/Virtual Channel Identifier) de la cabecera de la celda ATM.

Finalmente, MPLS también se ha desarrollado como una solución IP sobre Frame Relay. En este caso, descrito en la RFC 3034, la etiqueta es el DLCI (Data Link Control Identifier) de la cabecera Frame Relay.

3.9.2.1. Beneficios de MPLS

La migración a IP está provocando profundos cambios en el sector de las telecomunicaciones y configura uno de los retos más importantes para los ISP, inmersos actualmente en un proceso de transformación de sus infraestructuras de cara a incorporar los beneficios de esta tecnología. MPLS nació con el fin de incorporar la velocidad de conmutación del nivel 2 al nivel 3; a través de la conmutación por etiqueta; pero actualmente esta ventaja no es percibida como el principal beneficio, ya que los gigarouters son capaces de realizar búsquedas de rutas en las tablas IP a suficiente velocidad como para soportar todo tipo de interfaces. Los beneficios que MPLS proporciona a las redes IP son: realizar ingeniería del tráfico o TE (Traffic Engineering), cursar tráfico con diferentes calidades de clases de servicio o CoS (Class of Service) o grados de calidad de servicio o QoS (Quality of Service), y crear redes privadas virtuales o VPN (Virtual Private Networks) basadas en IP.

La TE permite a los ISP mover parte del tráfico de datos, desde el camino más corto calculado

Denominación	Entidades de red asociadas
E-UTRAN Uu	eNB UE
X2	eNB eNB
S1-MME	eNB Red troncal EPC (MME)
S1-U	eNB Red troncal EPC (S-GW)

Cuadro 3.4: Interfaces de E-UTRAN

por los protocolos de encaminamiento, a otros caminos físicos menos congestionados o menos susceptibles a sufrir fallos. Es decir, se refiere al proceso de seleccionar los caminos que seguirá el flujo de datos con el fin de balancear la carga de tráfico entre todos los enlaces, routers y switches en la red; de modo que ninguno de estos recursos se encuentre infrautilizado o sobrecargado. La TE, descrita en la RFC 2702, se ha convertido en la principal aplicación de MPLS debido al crecimiento impredecible en la demanda de recursos de red.

Mediante MPLS, los ISP pueden soportar servicios diferenciados o DiffServ, como viene recogido en la RFC 3270. El modelo DiffServ define varios mecanismos para clasificar el tráfico en un pequeño número de CoS. Los usuarios de Internet demandan continuamente nuevas aplicaciones, teniendo los servicios actualmente soportados unos requerimientos de ancho de banda y de tolerancia a retrasos en la transmisión muy distintos y para satisfacer estas necesidades óptimamente, los ISP necesitan adoptar no sólo técnicas de ingeniería de tráfico, sino también de clasificación de dicho tráfico. De nuevo, MPLS ofrece a los ISP una gran flexibilidad en cuanto a los diferentes tipos de servicios que puede proporcionar a sus clientes.

Finalmente, MPLS ofrece también un mecanismo sencillo y flexible para crear VPN. Una VPN simula la operación de una WAN (Wide Area Network) privada sobre la Internet pública. Para ofrecer un servicio de VPN viable a sus clientes, un ISP debe solventar los problemas de seguridad de los datos y soportar el uso de direcciones IP privadas no únicas dentro de la VPN. Puesto que MPLS permite la creación de circuitos virtuales o túneles a lo largo de una red IP, es lógico que los ISP utilicen MPLS como una forma de aislar el tráfico. No obstante, MPLS no tiene en estos momentos ningún mecanismo para proteger la seguridad en las comunicaciones, por lo que el ISP deberá conseguirla mediante cortafuegos y algún protocolo de encriptación tipo IPsec. Existen varias alternativas para implementar VPNs mediante MPLS, pero la mayoría se basan en la RFC 2547.

3.10. Interfaces de E-UTRAN

Ya se ha introducido en el capítulo anterior la reciente tecnología LTE (UMTS Long Term Evolution), explicando su composición a nivel general y, en particular, la arquitectura de su equivalente E-UTRAN (Evolved UTRAN), detallando al respecto las entidades de red características, los eNB (evolved Node B).

Como apoyo a dicha descripción, se han introducido ya a nivel general las diferentes interfaces que completan el sistema, recogidas en el cuadro 3.4. El propósito de este apartado es describirlas en mayor profundidad.

3.10.1. 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 3.22 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 contraposición 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.

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*).



Figura 3.22: Ilustración de los mecanismos de transferencia de información en la interfaz radio

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.

3.10.2. 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) 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 3.23 se ilustra dicho control del plano de usuario por parte de la entidad MME.
- ⇒ 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

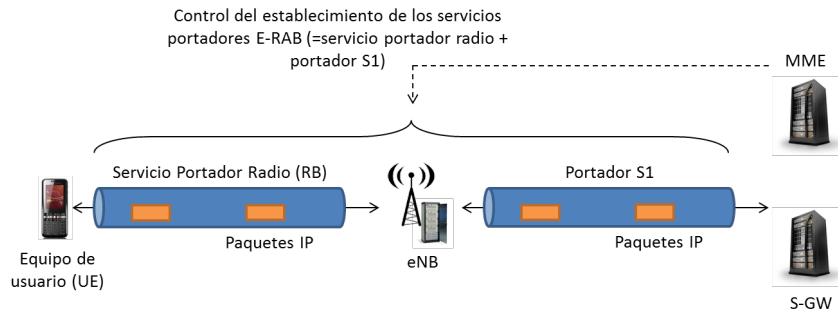


Figura 3.23: Control de los servicios portadores radio y S1 a través de la interfaz S1-MME

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.

- ⇒ Procedimiento de aviso (*Paging*). 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, cuando 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 Stratum*).

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.
- ⇒ 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.

3.10.3. 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.

3.11. 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 3.24 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 con objeto de comparar ambas arquitecturas.

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 25.40. En esta

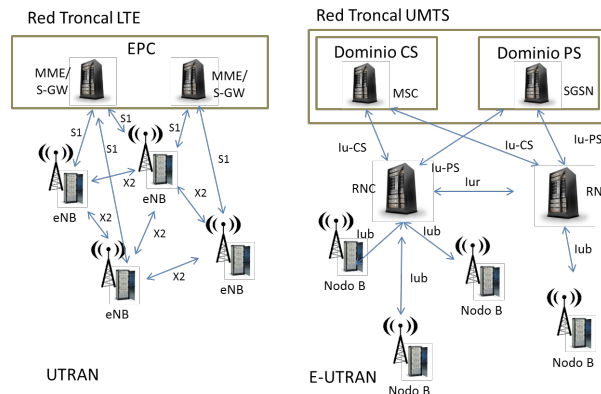


Figura 3.24: Comparativa de la arquitectura de red de E-UTRAN y UTRAN

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 Iub 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 sí mediante interfaces específicas como la interfaz Iur 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 Iu-PS, entre RNCs y los nodos SGSNs del dominio de paquetes, y Iu-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 3.24 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 sí 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 procesamiento 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. 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 procesamiento 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).

3.12. Conclusiones

Queda patente a la luz de la arquitectura presentada, que a diferencia de las redes GSM, en las redes UTRAN hay diferentes recursos cuya capacidad es gestionada de forma independiente:

- ⇒ Capacidad de procesamiento banda base.
- ⇒ Capacidad de spreading code WCDMA disponibles en las portadoras radio.
- ⇒ Capacidad en la interfaz Iub. Deben distinguirse flujos garantizados, como los ATM a tasa binaria constante y tiempo real (CBR RT) o los MPLS CRB RT – usados en servicios de voz –, de los no garantizados, como los ATM o MPLS VBR NRT –empleados en servicios de datos.

En el caso de MPLS sobre IP / Ethernet, debe considerarse el factor de agregación de flujos Ethernet en los sucesivos enlaces entre el nodo y la RNC.

El modelo propuesto para redes UTRAN puede ser extendido fácilmente a redes LTE con ligeras modificaciones. Por un lado, son diferentes los servicios ofrecidos por LTE, especialmente en el

ámbito de los datos a alta tasa binaria. Siguen existiendo tarjetas de procesado banda base, aunque con diferente cantidad de channel elements. En las portadoras ya no se utiliza WCDMA, sino OFDMA, dividiendo la capacidad en múltiples portadoras. Seguirá pudiéndose aplicar Kaufman Roberts, pero sin usar el concepto de spreading code. Y el e-Node ya no tiene enlace con la RNC, pero sigue teniendo un enlace con la red core, cuya capacidad sigue siendo un recurso escaso.

Capítulo 4

Modelado general KR sobre consumo de recursos

4.1. Fórmula de Erlang-B

4.1.1. Marco del modelo

Antes de la introducción de la moderna tecnología de redes de alta velocidad o banda ancha, la red telefónica conmutada acomodaba las llamadas telefónicas en circuitos físicos, de forma que para una conexión entre dos puntos unidos por una secuencia de líneas telefónicas se disponía en cada línea de un número de circuitos C que fijaba el máximo de conexiones telefónicas para una determinada línea de la red. Así, si estos circuitos estaban ocupados, sería rechazada una llamada que intentara establecer una conexión que usara esta línea. La fórmula de Erlang proporciona la probabilidad B de que al intentar establecer una llamada telefónica, ésta sea rechazada por estar ocupados todos los circuitos existentes en la línea que ha de utilizar esta llamada.

La probabilidad B en cada línea de una red telefónica es un parámetro de rendimiento básico. Si en una red de comunicaciones se fija un valor máximo B_m para este parámetro, se está definiendo un parámetro básico de la calidad de servicio que ofrece la red. Partiendo del conjunto de usuarios potenciales de una red de comunicaciones y del uso que necesitan hacer de la misma, el parámetro B_m determinará la capacidad de las líneas y por tanto será un criterio fundamental al diseñar la red; así mismo, un aumento en la capacidad de las líneas de la red conducirá a una disminución de la probabilidad de pérdida o bloqueo de las llamadas que los usuarios quieran realizar.

La fórmula de Erlang B se representa en la ecuación 4.1:

$$P_b = B(E, m) = \frac{\frac{E^m}{m!}}{\sum_{i=0}^m \frac{E^i}{i!}} \quad (4.1)$$

donde:

$\Rightarrow P_b$ es la probabilidad de bloqueo.

$\Rightarrow m$ es el número de recursos, en nuestro caso líneas telefónicas.

$\Rightarrow E = \lambda h$ es la cantidad total de tráfico en Erlangs. Donde λ es la tasa de llegada de nuevas llamadas y h la duración media de las mismas.

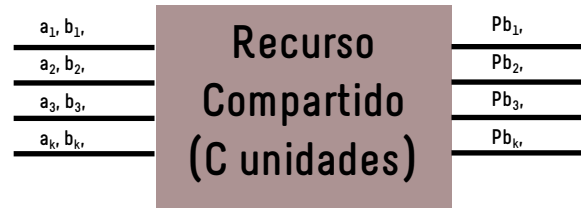


Figura 4.1: Marco del modelo de Kaufman-Roberts

Por ejemplo, si el escenario es tal que se tengan 3 llamadas por minuto, cada llamada tiene una duración de 2 minutos y se dispone de 7 líneas telefónicas en la central, es decir, $E = \lambda h = 6$, entonces aplicando ErlangB se obtiene que la probabilidad de bloqueo viene dada por 4.2.

$$P_b = B(6, 7) = \frac{6^7}{7!} \frac{1}{\sum_{i=0}^7 \frac{6^i}{i!}} = 0,18 \quad (4.2)$$

Es decir, que de cada 100 intentos de llamada, 18 no podrán realizarse por encontrarse el sistema totalmente ocupado.

El objetivo final de la teoría de Erlang es conseguir un grado de servicio (*Grade of Service*, GoS) o una calidad de servicio (*Quality of Service*, QoS) para la red que se desee dimensionar. Si en el escenario planteado se considera que la calidad de servicio debe ser tal que sólo 1 llamada de cada 100 no pueda cursarse, es evidente que debe aumentarse el número de líneas de la central de conmutación.

4.1.2. Limitación de la Fórmula de Erlang B

Antes de la aparición de las redes UMTS, la red telefónica conmutada acomodaba las llamadas telefónicas en circuitos físicos: para una conexión entre dos puntos unidos por una secuencia de líneas telefónicas se disponía en cada línea de un número de circuitos C que fijaba el máximo de conexiones telefónicas para una determinada línea de la red. De esta manera, si estos circuitos estaban ocupados, sería rechazada una llamada que intentara establecer una conexión que usara esta línea. Sin embargo, uno de los pilares de diseño de UMTS ha sido la capacidad de ofrecer múltiples servicios: voz, videllamada y transferencia de datos. Cada servicio supone una demanda distinta de recursos (líneas físicas) que no se pueden modelar mediante Erlang B y es, por tanto, por lo que es necesaria una expresión más completa.

4.2. Modelo de Kaufman-Roberts

4.2.1. Marco del modelo

El modelo propuesto por J. S. Kaufman y Roberts se enmarca dentro del escenario representado en la figura .

- ⇒ Se tiene un sistema que va a ser compartido y demandado por los clientes, que consta de « C » unidades.
- ⇒ Se tienen « k » tipos de clientes (o flujos, tráfico...), cada uno de ellos con diferente demanda de recursos, b_i .

⇒ Un cliente cuya demanda no puede ser satisfecha se dice que está bloqueado y desaparece del sistema sin afectar al mismo.

⇒ La llegada de clientes sigue un proceso de Poisson estacionario de media λ , siendo este valor la tasa de llegadas al sistema.

Este escenario se asemeja perfectamente al caso planteado en esta Tesis, ya que una de las características que definen las redes UMTS es que soportan diferentes tipos de servicios, los cuales demandan diferente cantidad de recursos. Como sistema se tiene la capacidad del enlace, de manera que todos los servicios comparten evidentemente la transmisión.

La política que define cómo se comparte ese recurso determina el bloqueo en el sistema. Kaufman y Roberts desarrollaron un algoritmo recursivo para el caso en que se tiene una política de compartición de recursos total: un cliente que requiere «b» unidades de un recurso será bloqueado si y sólo si las unidades que están disponibles en el sistema son menores que «b». Con él, es posible determinar el bloqueo que sufrirá cada tipo de tráfico.

4.2.2. Descripción del algoritmo multiservicio

El sistema puede ser descrito mediante un vector n , de manera que n_k representa el número de conexiones de clase «k». El enlace puede tener diferentes estados, y aquellos que están permitidos pueden describirse como el conjunto:

$$\Omega = \left\{ n : \sum_{i=0}^{\kappa} n_i b_i \leq C \right\} \quad (4.3)$$

Por otra parte, puede considerarse también β_k como el conjunto de estados permitidos donde no pueden admitirse más demandas de tipo «k» sin bloquear el sistema.

$$\beta_k = \left\{ n : C - b_k < \sum_{i=1}^{\kappa} n_i b_i \leq C \right\} \quad (4.4)$$

Por ejemplo, supóngase un sistema con capacidad de $C=5$ recursos, que soporta dos servicios diferentes:

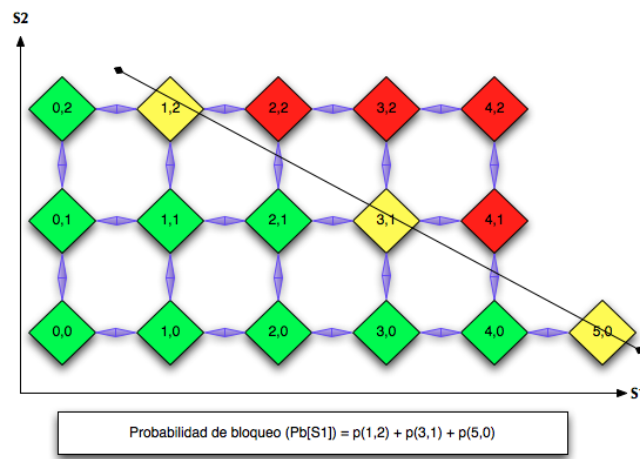
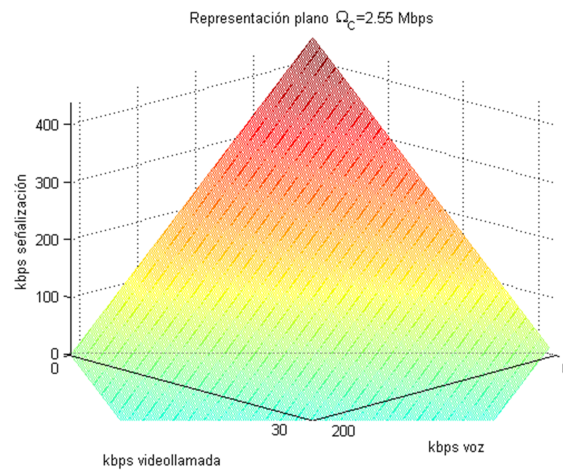
⇒ Servicio S_1 , que consume 1 recurso.

⇒ Servicio S_2 , que consume 2 recursos.

Como el servicio S_1 requiere de 1 recurso el máximo nº de usuarios demandantes del mismo se determina bajo el supuesto de que en el sistema no hay ningún usuario del resto de servicios, de manera que como máximo el sistema podrá atender a 5. De la misma manera, no podrán ser atendidos más de 2 clientes demandantes del servicio S_2 .

Una vez determinados los límites de Ω , es posible representar cualquier otra combinación de consumos en la que coexisten usuarios demandantes de los dos servicios simultáneamente. Puede recurrirse a una matriz que recoja todas las posibles combinaciones de servicios y, a partir de la misma, es posible calcular la probabilidad de bloqueo de cada servicio mediante la suma de las posibles combinaciones que originarían bloqueo, en el conjunto β (ver figura 4.2). Por ejemplo, para el servicio S_1 , las combinaciones que lo bloquean serían:

⇒ 1 cliente del servicio S_1 y 2 clientes demandantes de S_2 en el sistema.

Figura 4.2: Probabilidad de bloqueo servicio S_1 Figura 4.3: Conjunto Ω para el caso de servicios de voz, videollamada y señalización.

\Rightarrow 3 clientes S_1 y uno S_2 .

\Rightarrow 5 clientes S_1 y ninguno S_2 .

Así, la recta que une estos puntos representa el límite del conjunto Ω .

Aplicación en la Tesis

En el desarrollo de esta Tesis se calcula en primer lugar la accesibilidad de circuitos considerando 3 flujos de datos diferentes: voz, videollamada y señalización. Los cálculos pertinentes muestran que la demanda de recursos es, respectivamente, 12.2 kbps, 80 kbps, y 5.8 kbps, mientras que la capacidad estudiada es de $C=2.55$ Mbps. De esta manera, el conjunto Ω limitante puede representarse como un plano tridimensional, tal y como ilustra la figura 4.3.

4.2.3. Ecuaciones

Definiendo una función de densidad de probabilidad de ocupación total G , Kaufman demuestra que en el caso de una política de compartición completa, la probabilidad de bloqueo puede

determinarse relacionando los conjuntos Ω y β (ecs. 4.3 y 4.4) según la ecuación 4.5:

$$P_{B,i} = \frac{G(\beta_k)}{G(\Omega)} \quad (4.5)$$

En su procedimiento particular para el cálculo de estas funciones, Kaufman plantea el cálculo previo de las distribuciones de tráfico individuales, de manera que el bloqueo específico de cada clase de tráfico vendrá dado por la ecuación 4.6.

$$P_{B,i} = \sum_{i=0}^{b_i-1} q(C-i) \quad (4.6)$$

La fórmula en la ecuación 4.6 obtiene la probabilidad de bloqueo a partir de su distribución $q(j)$. Ésta se define como (ec. 4.7):

$$q(j) = \begin{cases} \frac{1}{j} \sum_{i=1}^k a_i b_i q(j-b_i) & \text{si } j > 0 \\ 0 & \text{si } j < 0 \end{cases} \quad (4.7)$$

donde, C es el número de recursos, k es el número de clases de servicio, a_i es la carga de tráfico ofrecida y b_i es la capacidad requerida por cada servicio.

Para el caso donde $j > 0$, se tienen los valores $q(j)$ expresados en función de $q(0)$, valor de normalización.

$$q(j) = \frac{1}{j} \sum_{i=1}^k a_i b_i q(j-b_i) \quad (4.8)$$

Finalmente, el cálculo de $q(0)$ es trivial, ya que $q(j)$ es una distribución, luego deberá satisfacer la condición (4.9).

$$1 = \sum_{i=0}^C q(j) \quad (4.9)$$

A partir de esta última relación es sencillo obtener los diferentes valores de $q(j)$.

4.2.4. Aspectos de la implementación

El algoritmo así descrito plantea obtener de forma recursiva los diferentes elementos de la distribución para obtener posteriormente obtener las probabilidades de bloqueo de cada clase de tráfico. En el sentido estricto, una función recursiva es aquella que para resolver el problema principal, se basa en las soluciones de sus problemas siguientes, es decir, la función se llama a si misma, reduciéndose la talla del problema, hasta llegar a unos casos tan simples que resulta trivial resolverlos. En este caso, podría reducirse la talla del problema en torno a dos parámetros: la capacidad y los recursos que consume cada clase de tráfico.

Es importante saber que aunque esta función devuelve los valores perfectamente, no es un algoritmo eficiente. Esto se debe a que un programa recursivo resuelve los subproblemas más veces de las necesarias, y en consecuencia, lo hace ineficiente. Para un vector de consumos de 3 componentes, siendo 1, 2, y 3 respectivamente, y una capacidad global del sistema de 6, ya se puede observar que el algoritmo internamente tiene que resolver sus subproblemas varias veces, y esto cuanto mayor es la capacidad ocurre con mayor frecuencia. Como consecuencia, colapsa

el computador. Por ello es necesario la optimización de este algoritmo convirtiendo el algoritmo recursivo a un algoritmo dinámico.

La programación dinámica es un procedimiento que intenta resolver problemas disminuyendo su coste computacional, pero aumentando su coste espacial. Básicamente, deben memorizarse todos los resultados de los subproblemas que se resuelvan, para que cuando el programa necesite volver a calcularlos éste sólo tenga que buscar su resultado, típicamente en una matriz. Para ello, se empieza almacenando en un vector el resultado correspondiente al caso más básico, cuando la capacidad es igual a cero, en la siguiente posición almacena el valor correspondiente a cuando la capacidad es igual a uno, y así sucesivamente. La implementación correspondiente en Matlab puede consultarse en los apartados A.3.5 y A.3.6.

4.2.5. Algoritmo de Kaufman revisado

Una de las limitaciones que presenta el algoritmo de Kaufman descrito anteriormente es que en algunas ocasiones se trabaja con números muy grandes que provocan un desbordamiento numérico en el computador, y que hacen por tanto imposible su cálculo. Este escenario se presenta fundamentalmente en 3 ocasiones:

- ⇒ Cuando la capacidad del sistema es elevada.
- ⇒ Cuando la intensidad de tráfico es elevada.
- ⇒ Cuando se tienen muchas clases de tráfico.

Una solución a este problema es introducir un esquema de normalización de resultados, por un valor suficientemente grande (e_{max}) que permita arrastrar cantidades grandes durante la recursión. Este procedimiento le añade al algoritmo de Kaufman una complejidad computacional del orden de $O(C^{2K}/2^k)$, de manera que se piensa en otras soluciones alternativas.

4.2.6. Mejora del algoritmo de Kaufman revisado basado en la FFT

Existen otros esquemas de pre-procesado que permiten no sólo trabajar con números grandes evitando el desbordamiento, sino también aligerar la carga algorítmica. Se presenta en este apartado un procedimiento que toma como base el algoritmo desarrollado por Kaufman.

Partiendo de la misma red de recursos compartidos, cuando la capacidad ocupada por el tráfico de clase «k» es $n_k b_k$, su probabilidad será cero si no es un múltiplo entero de b_k , como es lógico. En caso afirmativo, su valor está cuantificado, y Kaufman demostró que la función de distribución individual puede determinarse también por la expresión:

$$p_k(i) = \begin{cases} a_k^{n_k}/n_k! & i = n_k b_k \\ 0 & \text{resto} \end{cases} \quad (4.10)$$

El algoritmo implementado consta de los siguientes pasos:

1. A partir de la ecuación 4.10, es posible obtener de forma recursiva sucesivos elementos de la distribución a partir de uno dado. Para evitar el desbordamiento, se calcula en primer lugar el valor mayor, de manera que los sucesivos valores a calcular sean todos menores que éste. Por otra parte, a partir de las ecuaciones que se verán en los últimos pasos del algoritmo, se conoce cuál es el mayor valor que será necesario calcular. Si se conoce cuál es

el mayor valor que es capaz de manejar el computador sin desbordamiento («E»), basta con asegurarse de que los cálculos en este primer paso no excedan un determinado valor e_{max} , para, en posteriores cálculos no tener problemas de *overflow*. Puede demostrarse que el valor e_{max} que es necesario asignar al máximo de las funciones de densidad vendrá dado por la expresión:

$$e_{max} = \frac{1}{C} \sqrt[k]{\frac{E}{C^2}} \quad (4.11)$$

2. Determinación de la distribución de probabilidad de cada clase de tráfico. A partir de la ecuación 4.10, pueden distinguirse dos posibles formas para la función de distribución: convexa o creciente. De acuerdo con el razonamiento anterior, en este paso se implementa el cálculo de las densidades de probabilidad con las expresiones recogidas en 4.12:

$$\begin{aligned} \text{Si } n_{k,m\acute{a}x} > a_k & \\ p_k(n_k b_k) &= e_{max} \\ n_a &= \lfloor a_k \rfloor \\ p_k((n_k + 1) b_k) &= p(n_k b_k) a_k / (n_k + 1) \\ n_a &\leq n_k \leq n_{k,m\acute{a}x} \\ p_k((n_k - 1) b_k) &= p(n_k b_k) n_k / a_k \\ \text{Si } n_{k,m\acute{a}x} < a_k & \\ p_k(n_{k,m\acute{a}x} b_k) &= e_{m\acute{a}x} \\ n_{k,m\acute{a}x} &= \lfloor C / b_k \rfloor \\ p_k(n_k b_k) &= p_k((n_k + 1) b_k) (n_k + 1) / a_k \\ 0 &\leq n_k \leq n_{k,m\acute{a}x} \end{aligned} \quad (4.12)$$

3. Se define la función G de distribución de probabilidad de ocupación total como convolución de las distribuciones anteriores. Recurriendo a teoría de procesado de señal, este cálculo puede efectuarse utilizando productos sucesivos de transformadas de Fourier (ec. 4.13). Es en este paso donde los cálculos alcanzarían sus valores máximos.

$$G = p_1 \otimes p_2 \otimes \dots \otimes p_k = IFFT \left(\prod_{i=1}^K FFT(p_i) \right) \quad (4.13)$$

4. Finalmente, el cálculo de cada distribución de probabilidad vendrá determinado por la ecuación 4.5 vista en el apartado anterior:

$$P_{B,i} = \frac{G(\beta_k)}{G(\Omega)} = \frac{\sum_{i=C-b_k+1}^C G(i)}{\sum_{i=0}^C G(i)} \quad (4.14)$$

Este algoritmo presenta una complejidad algorítmica del orden de $O(\log_2 C (KC/2))$, y por tanto permite el cálculo de bloqueo para números elevados sin perjuicio de desbordamiento y en un periodo de tiempo razonable. El lector puede consultar la implementación del mismo en Matlab en el apartado A.3.7.

4.3. Comparación del rendimiento de los algoritmos

La implementación del algoritmo de Kaufman-Roberts para redes multiservicio se ha venido realizando de manera secuencial, tal y como ha sido descrito a lo largo del presente capítulo. Los detalles del material empleado en las simulaciones se pueden encontrar en el Apéndice H. Las investigaciones en torno a este problema se han realizado de la siguiente manera:

1. Se implementa el algoritmo original propuesto por Kaufman-Roberts [90] para estudiar el comportamiento de la red en cuanto a la compartición del canal Iub. Inmediatamente se observa que los tiempos de simulación son muy elevados, y además, cuando la capacidad del sistema crece, el sistema se desborda en los cálculos más críticos.
2. Se opta por implementar entonces el algoritmo revisado, observando que el tiempo de computación es ligeramente inferior para capacidades pequeñas, pero que crece exponencialmente según se incrementa la capacidad del sistema. Además, sigue manteniéndose el mismo problema de *overflow* para este caso.
3. Se desarrolla finalmente el algoritmo basado en la FFT, que además de evitar explícitamente el desbordamiento consigue reducir los tiempos de ejecución de forma muy significativa. Recurriendo a este algoritmo como base, se implementarán los estudios posteriores.

En la figura 4.4 se presenta el crecimiento del tiempo de ejecución del algoritmo para las 3 implementaciones, donde se hace variar la capacidad del sistema para ver cómo responde el algoritmo. Esta variación se lleva a cabo de forma incremental para verificar en qué punto los algoritmos presentan o no problemas de desbordamiento.

Cabe destacar que estos tiempos de cálculo corresponden al estudio de una única muestra en un instante dado, mientras que los estudios que se desarrollarán a lo largo de esta Tesis corresponden a un día completo (1 muestra por minuto, 1440 muestras en total, 1440 ejecuciones diferentes en batch de este algoritmo). Por otra parte, en los casos de desbordamiento el consumo de recursos computacionales es el mismo y por tanto también lo es el tiempo empleado en su iteración, aunque el resultado final mostrado sea «NaN» (*Not a Number*).

En cuanto a tiempos de ejecución se refiere, puede verse que cuando la capacidad del sistema es relativamente pequeña, los 3 algoritmos se comportan de forma muy parecida, si bien el algoritmo original propuesto por Kaufman-Roberts presenta un comportamiento ligeramente menos eficiente. Como puede comprobarse en la gráfica 4.4, según aumenta la capacidad del sistema la tendencia del algoritmo original y el basado en la FFT presentan un comportamiento más lineal, mientras que el algoritmo revisado se desmarca de los otros 2 y presenta un comportamiento exponencial.

En cualquier caso, el algoritmo basado en la FFT es el que presenta un mejor comportamiento en todas las situaciones, reduciendo en algo más de un 50 % los tiempos de ejecución respecto al algoritmo original.

4.4. Conclusiones

La heterogeneidad característica de los tráfico que soporta la red UMTS plantea un escenario en el que se tienen tráfico que en principio demandan individualmente diferente cantidad de recursos. El sistema, por otra parte, dispone de una cantidad de recursos limitados. Por tanto, todos los tráfico «compiten» al mismo tiempo por un mismo recurso.

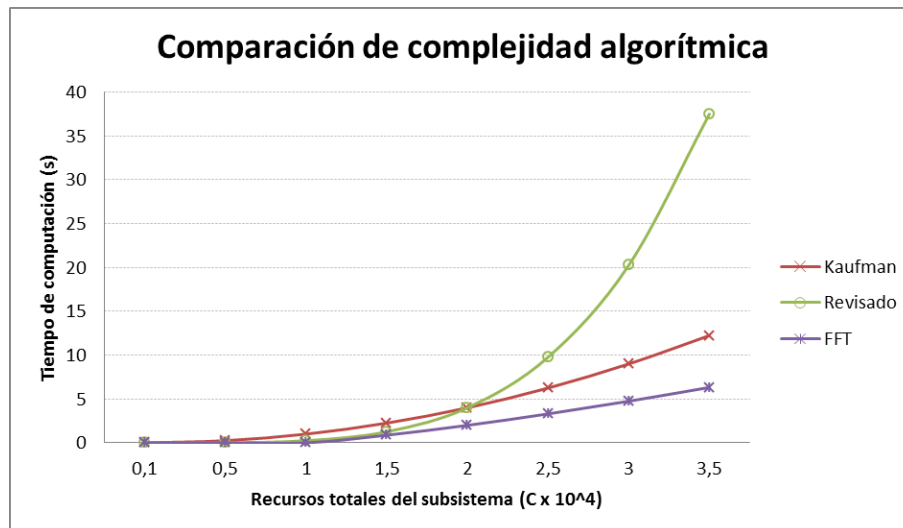


Figura 4.4: Comparación de la complejidad algorítmica de las diferentes implementaciones de Kaufman-Roberts (tiempos por iteración/muestra)

Esta situación ha llevado al autor a pensar en la aplicación directa de un algoritmo que permita analizar el bloqueo de cada uno de los servicios. La implementación de este algoritmo y su posterior validación numérica han dado pie a que durante la Tesis se desarrolle una investigación más exhaustiva en este campo. Las conclusiones que al autor de esta Tesis plantea en relación con este apartado son las siguientes:

- ⇒ La naturaleza de la red UTRAN como red mutiservicio hace evidente que no es posible utilizar el modelo de Erlang-B para el modelado de la misma. El modelo de Kaufman-Roberts se ajusta de una forma muy precisa la situación de bloqueo multiservicio característica de una red UTRAN, como así lo atestigua el ensayo numérico de diferentes escenarios reales y su contraste con las estadísticas reales.
- ⇒ El proceso en su conjunto requiere, para su validación, la ejecución iterativa del mismo algoritmo abstracto, particularizado a diferentes situaciones (tipos de tráfico rígido, elástico, enlace ascendente y descendente). Por tanto, es imprescindible encontrar un algoritmo fiable y con una complejidad algorítmica baja para que la herramienta sea manejable.
- ⇒ El algoritmo primitivo planteado por Kaufman y Roberts en la década de los 80 funciona muy bien cuando se trabaja con un número de peticiones pequeña y con una capacidad de sistema limitada. La validación del algoritmo bajo cantidades de mayor magnitud da como resultado el desbordamiento numérico y, dada su alta complejidad algorítmica, una ejecución muy lenta, por lo que en estos supuestos no es recomendable ni adecuada la implementación de este algoritmo.
- ⇒ Existen otros planteamientos del algoritmo de Kaufman-Roberts mejorados que palian con éxito las carencias del algoritmo original. En particular, y como base para el desarrollo de la Tesis, se han consultado diversas fuentes y se ha implementado una sucesión de algoritmos que han ido mejorando la calidad de las simulaciones. Un algoritmo concreto basado en la FFT ha resultado ser válido y fiable, evita el desbordamiento numérico y al mismo tiempo posee una complejidad algorítmica mucho más sencilla, lo que se traduce en cálculos más

rápidos.

Capítulo 5

Caracterización de servicios en una red UTRAN

5.1. Introducción

Uno de los objetivos principales de UMTS es facilitar el acceso a una amplia gama de servicios, mayor que la previamente disponible en redes GSM/GPRS. La consecución de este objetivo se materializa mediante la formalización de una serie de mecanismos de provisión de servicio.

El punto de partida para posibilitar el soporte de múltiples servicios es la definición de una arquitectura de servicios portadores capaz de adaptarse a los requisitos de las distintas clases de tráfico previstas en UMTS. Este objetivo se concreta en la definición de los mecanismos de calidad de servicio QoS.

Un segundo conjunto de mecanismos encaminados a facilitar el soporte de servicios en UMTS son los encuadrados bajo el concepto VHE (*Virtual Home Environment*). Su objetivo es permitir la personalización y portabilidad de servicios, facilitando su provisión y desarrollo por terceros.

5.2. Servicios Portadores

Siguiendo un enfoque contrario a los sistemas 2G, la provisión de servicios en UMTS se aborda desde una perspectiva abierta. Así, en vez de especificar un repertorio cerrado de servicios finales, se apuesta por la definición de una arquitectura de servicios portadores robusta y flexible, capaz de adaptarse a los requisitos de QoS de las distintas aplicaciones a soportar.

La arquitectura de los servicios portadores UMTS, definida en la especificación TS.23.107, se basa en el modelo jerárquico mostrado en la figura 5.1. La calidad de un servicio final UMTS, tal como la percibe el usuario, viene determinada por la de los distintos tramos involucrados en el trayecto de la comunicación. En una primera aproximación, el trayecto se descompone en el servicio portador local, el servicio portador UMTS y el servicio portador externo. Las especificaciones del 3GPP se centran únicamente en el tramo correspondiente al servicio portador UMTS. De este modo, se evita imponer limitaciones innecesarias acerca de los terminales a emplear o las posibles redes destino.

El servicio portador UMTS se descompone en el servicio portador de acceso radio o *Radio Access Bearer* (RAB), y el servicio portador del núcleo de red. El primero abarca el trayecto comprendido

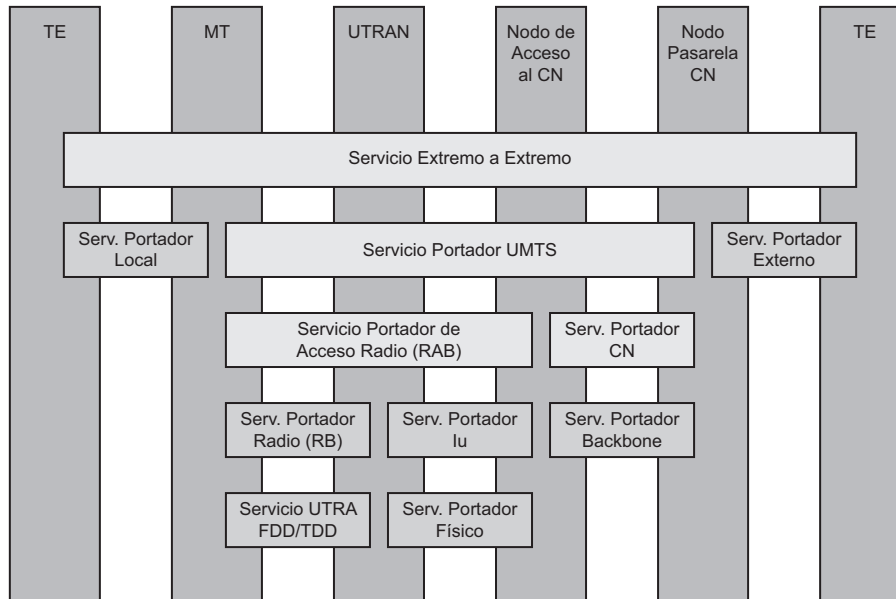


Figura 5.1: Arquitectura de servicios portadores UMTS

entre el móvil y el nodo de acceso al núcleo de red (un MSC o SGSN, según el caso). El servicio portador del núcleo de red, por su parte, abarca el tramo comprendido desde el nodo de acceso (MSC o SGSN) hasta el nodo pasarela (GMSC o GGSN) hacia la red destino de interés.

El RAB es responsable de garantizar los recursos necesarios para el intercambio de información entre el móvil y el núcleo de red. Es determinante en la provisión de servicios UMTS con distintos perfiles de QoS, ya que implica la utilización de recursos sobre la interfaz radio y la red de acceso, precisamente donde se dan las mayores limitaciones de capacidad. En función de los servicios a ofrecer por el operador, una red UMTS puede ofrecer distintos tipos de RAB.

Dentro de él se distinguen dos tramos, cada uno de ellos sustentado por el correspondiente servicio portador. Así, el servicio portador radio, RB (*Radio Bearer*) abarca el trayecto comprendido entre el móvil y la RNC, incluyendo todos los aspectos relativos al transporte de la información sobre la interfaz radio UTRA, así como las interfaces Iub. Dentro del RAB se incluyen también las facilidades para el transporte de información entre la red de acceso y el núcleo de red, esto es, sobre la interfaz Iu.

Asociado a cualquier servicio portador sobre un RAB se requiere de una conexión paralela encargada de la señalización asociada, denominada *Signalling Connection*, y, que al igual que el RAB, se distinguen dos tramos: por un lado el comprendido entre el móvil y la RNC, análogo al RB, y, por otro lado, un último tramo de señalización entre la RNC y el núcleo de red por el que se intercambian mensajes de señalización RANAP. En la figura 5.2 se muestra el esquema final del RAB y la señalización asociada.

La calidad de servicio portador del núcleo de red, por su parte, se apoya en la que le proporciona su «*backbone*» (de circuitos o paquetes). El soporte de QoS en este tramo presenta menos dificultades que en la red de acceso ya que, habitualmente, se dispone de capacidades de transmisión elevadas. El problema se reduce, básicamente, en aplicar un tratamiento prioritario a los servicios con requisitos de retardo más estrictos.

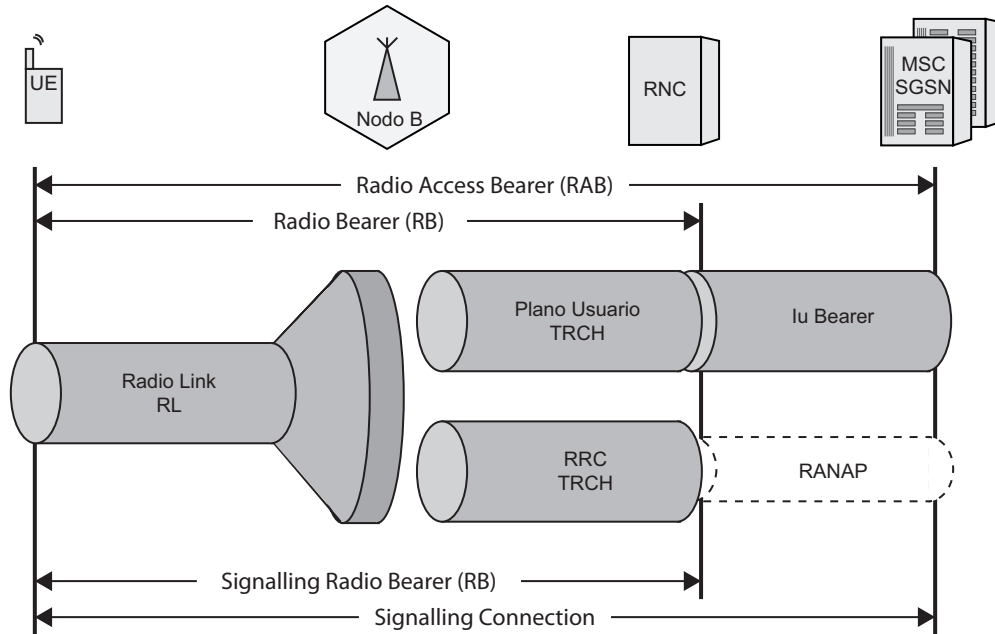


Figura 5.2: *Radio Access Bearer* (RAB) y señalización

5.3. Clases de Tráfico

Desde el punto de vista del usuario final podemos identificar cuatro grandes clases de tráfico. A continuación se describen detalladamente y en el cuadro 5.1 se resumen sus puntos más importantes:

5.3.1. Aplicaciones en tiempo real

5.3.1.1. Clase conversacional (Conversational class)

Dentro de esta clase se encuadran las comunicaciones de audio y vídeo en tiempo real entre personas. Este tipo de comunicaciones se caracteriza por exigir un retardo extremo a extremo muy reducido, con objeto de que los usuarios no pierdan la sensación de interactividad: el máximo retardo en la transmisión viene dado por la propia percepción humana de lo que es una conversación de audio o de vídeo.

En resumen, las cualidades básicas requeridas para este tipo de tráfico son: bajo retardo, bajo jitter (variación en el retardo), claridad razonable (calidad del codec) y ausencia de ecos. En el caso de aplicaciones multimedia, como la videoconferencia, se hace también necesario mantener una correcta temporización relativa entre los diferentes flujos, por ejemplo sincronización labial.

Esta clase es tolerante a errores, ya que la pérdida de un paquete o su corrupción (en tiempos de en torno a 20 ms) no es perceptible por el usuario final. El grado de protección frente a errores que se necesita va a depender en algunas aplicaciones. Algunos algoritmos de compresión dividen la información en varios conjuntos con tratamientos distintos. Por ejemplo, con videoconferencia H.324, la información asociada con el movimiento labial es más crítica que la calidad del escenario de fondo, de manera que a la primera puede ofrecérsele una protección frente a errores más alta. Para el caso de la voz, el códec AMR definido para UMTS tiene 3 categorías diferenciadas.

Uno de los puntos claves, preliminar, para la negociación de la QoS es la elección del codec. Si los codecs empleados por los dos usuarios fuesen diferentes, entonces se haría necesario el empleo

de una pasarela de conversión de información que introduciría retardo y por tanto, un descenso en la calidad percibida. La elección del codec debe ser compatible también con los recursos de red disponibles, donde intervendrán la RNC, el SGSN, y el GGSN.

Ejemplos de aplicaciones conversacionales son la telefonía y la videollamada. No obstante, con el uso de Internet y de aplicaciones multimedia, hay un buen número de nuevas aplicaciones que pueden utilizar este esquema, por ejemplo Voz sobre IP (VoIP) y aplicaciones de videoconferencia.

5.3.1.2. *Streaming Class:*

Aplicaciones de descarga de contenidos multimedia (audio y videoclips) para reproducción online, con una sensación que, sin serlo, se aproxima a la de tiempo real. El hecho de que la transferencia de información sea unidireccional permite retrasar el instante de inicio de la reproducción usando *buffers* relativamente grandes en el extremo receptor para absorber las fluctuaciones del retardo o *jitter*. Ello permite relajar significativamente los requisitos del retardo frente a los servicios conversacionales.

5.3.2. Aplicaciones NO tiempo real

5.3.2.1. *Interactive class:*

Aplicaciones de acceso remoto a la información en la modalidad *online*, donde el usuario envía peticiones hacia el equipo remoto esperando que éste le devuelva las respuestas en tiempo razonablemente reducido. Este esquema se aplica cuando el usuario final, que puede ser un usuario o una máquina demanda datos de un equipo remoto (por ejemplo, de un servidor). Ejemplos de interacción con un usuario pueden ser: navegación web, peticiones a bases de datos, acceso remoto a ordenadores y acceso a servidores. Como ejemplo de interacción con máquinas cabe citar los equipos de telemedida. Uno de los atributos clave dentro de este esquema es el round-trip delay time (RTT), es decir, el tiempo de ida y vuelta de un paquete pasando por el destino. Otra característica es que el contenido de los paquetes se transfiere de manera transparente, con una baja tasa de error.

5.3.2.2. *Background class:*

Este esquema aplica cuando el usuario final (típicamente un ordenador) envía y recibe ficheros en segundo plano. Como ejemplos pueden citarse el reparto de correo electrónico, SMS, descarga de bases de datos o la recepción de datos de telemetría. El tráfico de background es uno de los formatos más típicos de telecomunicaciones en los que a nivel general se caracterizan por el hecho de que el destino no está esperando la llegada de datos en un tiempo determinado, por lo que se trata de un esquema más o menos «insensible» a los tiempos de entrega. Otra característica es que el contenido de los paquetes se transfiere de forma transparente, con una baja tasa de error.

La clasificación anterior de aplicaciones en tiempo real y no tiempo real da origen a lo que se llaman flujos garantizados y no garantizados. Dicho flujo se compone esencialmente de la pareja RAB más SRB. En el apartado siguiente se describen las características principales de ambos tipos.

Los parámetros más importantes en cuanto a niveles de QoS se resumen en el cuadro 5.2. El principal factor que distingue a las 4 clases de tráfico es cómo es de sensible al retardo, por ejemplo la clase conversacional es la más sensible, mientras que la clase de background es la más insensible. Hay otros 3 parámetros más a tener en cuenta. El bitrate garantizado define el mínimo bitrate

Clase de tráfico	Características	Ejemplos
Conversacional	<ul style="list-style-type: none"> ⇒ Patrón conversacional (bajo retardo, muy crítico). ⇒ Mantener la relación temporal entre las entidades de información del flujo. 	Voz
Streaming	<ul style="list-style-type: none"> ⇒ Mantener la relación temporal entre las entidades de información del flujo. 	Difusión de vídeo
Interactiva	<ul style="list-style-type: none"> ⇒ Patrón de petición/respuesta. ⇒ Mantener la integridad del contenido 	Navegar por www
Segundo plano	<ul style="list-style-type: none"> ⇒ No hay un límite estricto al tiempo de transferencia de datos. ⇒ Mantener la integridad del contenido 	Descarga de correo

Cuadro 5.1: Resumen tipos de tráfico en UMTS según QoS

	Conversacional	Streaming	Interactivo	Background
Retardo	80 ms -	250 ms -		
Bit rate garantizado	Más de 2 Mbps	Más de 2 Mbps		
Prioridad en el manejo de tráfico			1, 2, 3	
Prioridad de asignación/retención	1, 2, 3	1, 2, 3	1, 2, 3	1, 2, 3

Cuadro 5.2: Parámetros de QoS para tráficos en UMTS

en el RAB que la red UTRAN debe proveer, y puede usarse en el control de admisión y en la asignación de recursos. El tiempo de retardo define el 95-percentil del retardo que se requiere, y puede utilizarse para definir el modo de operación RLC y el n° de transmisiones.

Las clases de servicio, en lo que se refiere a QoS, no son de obligado cumplimiento para la inclusión de cualquier servicio de bajo retardo, en las redes no se ha hecho demasiado uso de las mismas en tal caso. Es posible soportar video en streaming o voz sobre IP conversacional desde el punto de vista de rendimiento end-to-end usando solamente la clase background.

La diferenciación entre clases de servicio se vuelve útil para la eficiencia de la red durante periodos de alta carga de tráfico cuando coexisten servicios caracterizados por tiempos de retardo diferentes. Si la red radio tiene conocimiento sobre los requisitos de retardo de las diferentes clases de tráfico sería capaz de priorizar los servicios de forma correcta mejorando así la eficiencia de uso de la red. En la figura 5.3 se muestra la ganancia de QoS en términos cualitativos. En el paso n° 2 se pueden introducir mejoras en la eficiencia sólo por el hecho de introducir unas pocas clases de priorización dentro de las clases interactivas y de background usando los parámetros de retención y asignación (ARP).

Como puede verse en el paso n° 3, la priorización de paquetes por sí sola no es suficiente

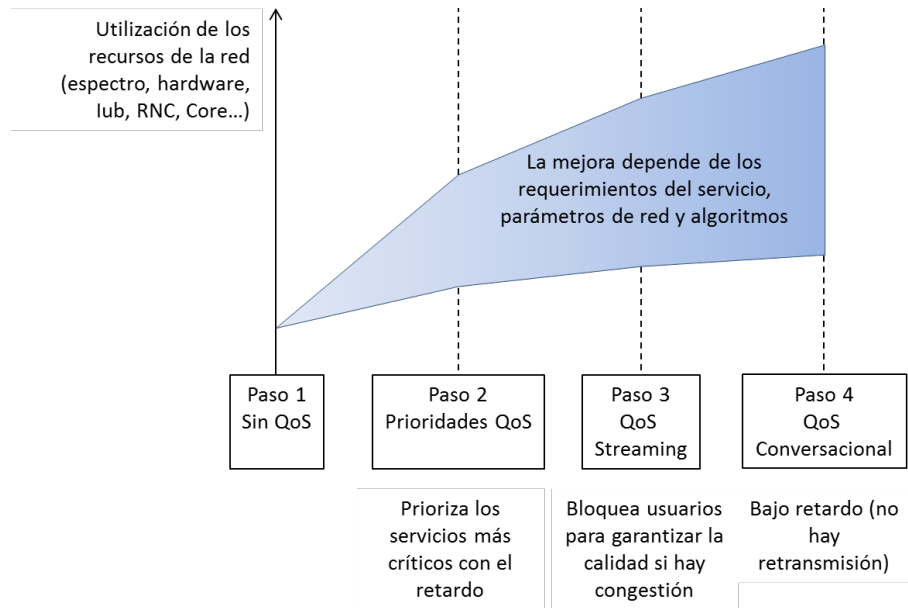


Figura 5.3: Ganancia cualitativa según la diferenciación en QoS

para proveer mejoras en la diferenciación de QoS completas. Usuarios dentro de la misma QoS y clase ARP compartirán la capacidad disponible. Si el nº de usuarios es demasiado alto, todos sufrirán una baja calidad. En tal caso sería más conveniente bloquear a unos pocos usuarios para garantizar la calidad de las conexiones ya existentes, como por ejemplo video en streaming. La red puede estimar la capacidad radio disponible y bloquear un usuario entrante si no hay recursos suficientes para proveerle del ancho de banda necesario sin sacrificar la calidad de las conexiones ya existentes.

El paso nº 4, es el que muestra más diferencia entre servicios con bitrate garantizado con diferentes requisitos de retardo. Si se conocen estos requisitos, el WCDMA RAN puede ubicar los parámetros de radio apropiados, como parámetros de retransmisión, para el nuevo cliente.

5.4. Atributos de Calidad de Servicio

la caracterización de los requisitos de QoS de los servicios portadores UMTS se concreta en base a un conjunto de atributos claramente definidos. La especificación proporciona también indicaciones acerca de los rangos de valores permitidos para cada atributo.

En la tabla 5.3 se muestran los atributos de QoS correspondientes al servicio portador RAB para cada una de las cuatro clases de tráfico UMTS. Nótese que no todos los atributos son aplicables a todas las clases de tráfico.

5.5. Mecanismos de Calidad de Servicio

En la especificación TS 23.107, se identifican de manera genérica las funciones de gestión de QoS a proporcionar en una red UMTS, así como su posible ubicación. En la tabla se indican las funciones de QoS definidas tanto en el plano de usuario como en el de control.

	Conversacional	Streaming	Interactivo	Background
Tasa de bit máxima	X	X	X	X
Entrega ordenada	X	X	X	X
Tamaño máximo SDU	X	X	X	X
Información de formato de SDU	X	X		
Ratio de SDUs erróneos	X	X	X	X
Ratio de error de bit residual	X	X	X	X
Entrega de SDUs erróneas	X	X	X	X
Retardo de transferencia	X	X		
Tasa de bit garantizada	X	X		
Prioridad de tráfico			X	
Prioridad de asignación/retención	X	X	X	X

Cuadro 5.3: Atributos de QoS del servicio portador RAB

Plano	Función	Descripción
Control	Gestor servicio	Coordinación de funciones del plano de control para establecer, modificar y mantener el servicio.
	Traducción	Convierte primitivas de control de servicio de UMTS a las de las redes externas.
	Control Admisión	Decide si acepta o no una conexión en función de los recursos disponibles.
	Control Suscripción	Comprueba si se cumplen los permisos administrativos necesarios para acceder al servicio.
Usuario	Correspondencia	Marca las unidades de datos en función de la calidad de servicio requerida.
	Clasificación	Distribuye las unidades de datos sobre los servicios portadores en curso según su calidad de servicio.
	Gestor de recursos	Gestiona la utilización de recursos según la calidad de servicio requerida (ej. planificación)
	Acondicionador tráfico	Adapta el tráfico al perfil de calidad de servicio negociado.

Cuadro 5.4: Funciones de gestión de QoS en UMTS

5.5.1. Servicios sobre flujos garantizados

El hecho de que un servicio viaje sobre un flujo de datos garantizado implica que se deberá asegurar la conexión sobre el Iub y, por tanto, se deberán reservar los recursos que sean precisos para lograr tal objetivo. Otra consecuencia es que cualquier conexión de menor prioridad podría llegar a ser excluida para liberar ancho de banda.

Cuando la tecnología empleada es ATM los flujos garantizados viajarán a través de caminos y canales virtuales con descriptores de tráfico de tipo CBR y conexiones de clase A o B.

Estos flujos requieren que se implementen mecanismos de control que monitoricen el estado de la red a fin de pronosticar si se puede o no iniciar una nueva conexión con los requerimientos mínimos necesarios. Normalmente este mecanismo se conoce como Control de Admisión y en las redes UMTS se implementa en el elemento RNC.

El hecho de que exista un mecanismo que decida si un servicio puede o no puede ser soportado con una calidad de servicio dado conduce directamente a la existencia del bloqueo de conexiones. Existirá, pues, una probabilidad de bloqueo en función del tráfico demandado y de la capacidad de transmisión disponible.

En la red bajo estudio los servicios que emplean flujos garantizados son los de voz, videollamada y señalización, independientemente del *codec* o conexión empleados.

5.5.2. Servicios sobre flujos no garantizados

En contraposición a los flujos garantizados, los no garantizados no son más que servicios que aprovechan el excedente de ancho de banda que no utilizan los primeros.

Este tipo de flujo se utilizará para transporte de información con menor requerimiento de calidad de servicio y, por tanto, se empleará con conexiones de datos.

UMTS dispone de varios tipos. Por un lado, existe lo que se conoce como tráfico de datos R99, el cual admite velocidades fijas de transmisión de 64kbps, 128kbps y 384kbps y tráfico de datos HSPA (denominado HSDPA cuando es tráfico *downlink* y HSUPA cuando es tráfico *uplink*), cuya capacidad de transmisión es completamente flexible y ajustará la velocidad al número de usuarios y ancho de banda sobrante (figura 1.1).

Mientras que en R99 sí puede encontrarse el concepto de bloqueo, ya que, por ejemplo, sobre un único enlace de 2Mbps se pueden cursar un máximo de 5 conexiones a 384kbps, 15 conexiones a 128kbps ó 31 conexiones a 64kbps, en HSPA no existirá bloqueo en los recursos de transmisión, ya que sobre esos 2Mbps se podrían tener cualquier número de conexiones simultáneas repartiéndose equitativamente el ancho de banda. Existirán límites al número máximo de conexiones HSPA simultáneas pero que vendrán marcadas por otro tipo de recursos distintos al ancho de banda de la transmisión.

5.6. Conclusiones

Como se ha visto, la red UTRAN da cabida a servicios que demandan una capacidad de red garantizada durante la conexión (servicios rígidos), mientras que hay otros servicios que admiten una capacidad variable durante la conexión (servicios elásticos). Por otra parte, hay servicios que requieren tiempo real, como la voz y la videotelefonía, y hay servicios de bajo consumo en términos de capacidad de procesamiento y tasa binaria de transferencia – como la voz –, mientras que la demanda de otros es muy elevada – como HSDPA y HSUPA. Ante este panorama, se puede comprender que

se hace necesario un modelado multiservicio, que contemple estos diferentes requerimientos. Y es necesario también que este modelado multiservicio se traslade a cada uno de los recursos por los que de forma consecutiva habrá de progresar la conexión entre el móvil (User Equipment) y la RNC.

Capítulo 6

Caracterización de recursos en una red UTRAN

6.1. Introducción

Tal y como se ha visto anteriormente, la red bajo estudio sirve de soporte para diferentes servicios. Los algoritmos de control de admisión (CAC) que se implementan en la RNC sirven para gestionar la calidad del servicio en redes multiservicio, de manera que si bien el tráfico alcanza cotas altas y momentos de elevada congestión, los recursos disponibles son limitados y en ocasiones escasos, tal y como demuestran los KPI.

Si bien la caracterización de los servicios ha sido objeto de estudio en el capítulo anterior, en el presente se explican detalladamente los recursos de que se dispone en este tipo de escenarios.

El esquema general de cómo se gestionan los recursos se muestra en la figura 6.1.

En capítulos anteriores se ha descrito la teoría de Erlang y el marco en el que puede aplicarse, si bien se ha introducido el modelado de Kaufman-Roberts para redes multiservicio, en la que existe un recurso escaso y varios usuarios con diferentes demandas de ese servicio compiten por el mismo. En WCDMA, la fórmula de Erlang se puede utilizar casi directamente en redes de conmutación de circuitos (CS), puesto que durante una llamada los recursos están dedicados exclusivamente al usuario que la realiza. La diferencia más importante entre las tecnologías WCDMA y FDMA/TDMA es la gran cantidad de recursos que están definidos, tal y como se ilustra en la figura 6.2. Tal y como indica la figura, los recursos deben definirse y planificarse para cada componente de la red WCDMA.

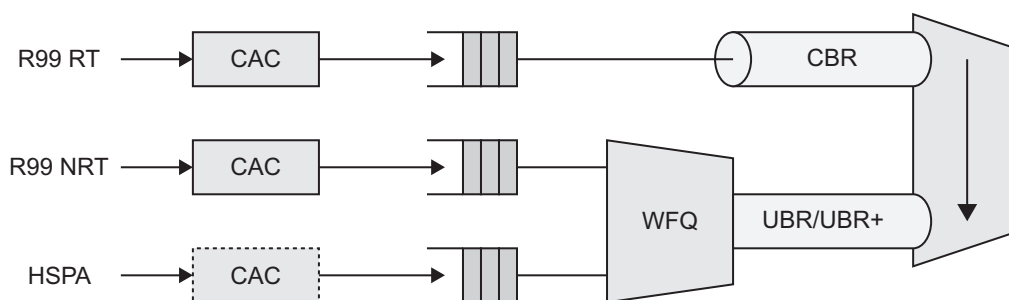


Figura 6.1: Esquema de la gestión de tráfico

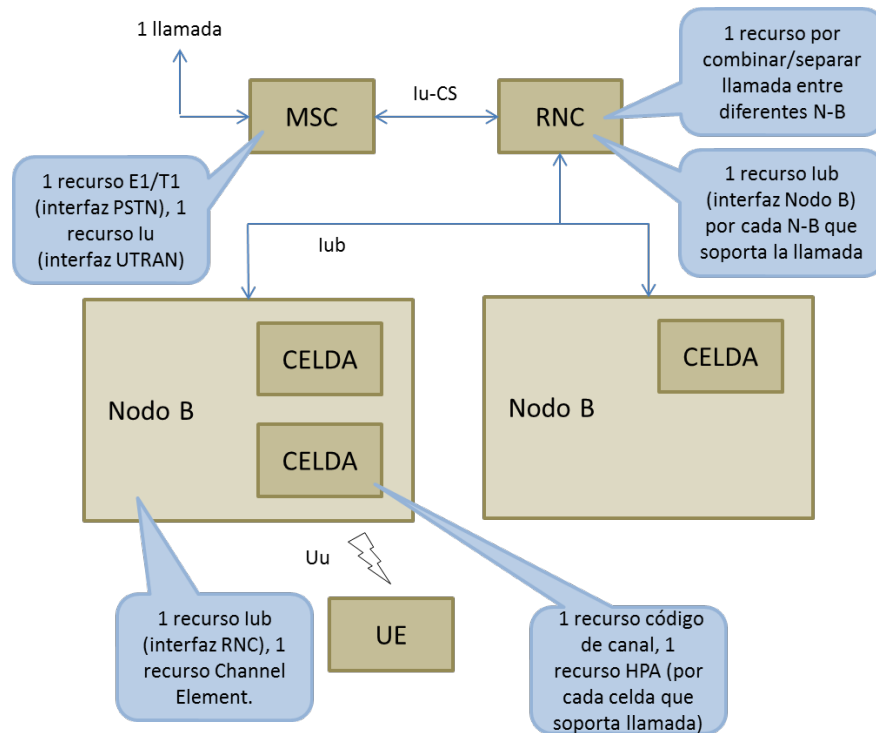


Figura 6.2: Definición de recursos en un sistema WCDMA

- ⇒ En la MSC, los recursos se asignan para la duración de la llamada; la teoría de Erlang se puede aplicar de forma directa. Las suposiciones de tráfico que son necesarias están referidas al tráfico de los usuarios, el cual es equivalente al tráfico de la MSC. Todos los tipos de tráfico se incluyen en este punto, a pesar de ser un punto de terminación.
- ⇒ En la RNC, los recursos deben definirse para las interfaces Iu e Iub. La teoría de Erlang puede aplicarse directamente para los recursos de la interfaz Iu y para los recursos comunes, mostrados en la figura como una función de combinación/separación. Para los recursos de la interfaz Iub, también podría aplicarse la teoría de Erlang, pero el tráfico habría de multiplicarse por un factor representativo del soft-handover. En la RNC pueden definirse recursos adicionales, pero serían muy variables en función de la implementación que se considerase. Para una arquitectura genérica, los recursos podrían dividirse en comunes (uno por llamada) y dedicados (uno por Iub).
- ⇒ En el nodo B, los principales recursos a dimensionar son: las portadoras (una o múltiples frecuencias), channel elements, códigos de canalización y potencia. Además, el dimensionamiento debe realizarse de tal manera que el crecimiento de ruido en el uplink tenga el menor impacto posible en la cobertura de la celda. Los channel elements son el recurso más fácil de dimensionar, puesto que se necesita uno por llamada y por nodo B, a pesar del porcentaje de soft-handover. El dimensionamiento correcto de channel elements es, no obstante, importante para garantizar que no habrá bloqueos en el uplink o en el downlink. Los códigos de canalización deben dimensionarse de acuerdo con el soft-handover, se requiere un código por cada asociación de handover. La potencia necesaria para cada enlace depende de las condiciones de radio y del estado de handover. El dimensionamiento debe hacerse a través de simulaciones

Cliente	Voz	Datos	Corporate
Población objetivo	1000000	1000000	100000
Objetivo de penetración	40 %	10 %	20 %
MoU por cliente	1000	1000	2000
MB mensuales por cliente	10	100	200

Cuadro 6.1: Ejemplo de supuestos de tráfico utilizados durante una planificación inicial

para un conjunto dado de supuestos. Los códigos de canalización y los recursos de potencia limitan fundamentalmente el downlink, mientras que el ruido en el uplink limita el tráfico del uplink.

6.1.1. Requisitos de capacidad

En las redes de conmutación de circuitos (en adelante, dominio de circuitos o CS), los requisitos del tráfico se definen en términos de Erlangs, durante la hora de máxima ocupación. No obstante, no siempre coincide con las consideraciones de marketing. Cada definición que se da, no es práctica para la planificación de red porque no muestra la distribución de usuarios dentro de una determinada área.

La aproximación está basada en un requerimiento de tráfico genérico, ajustado por los requisitos de la hora de máxima ocupación, teniendo en cuenta la distribución geográfica, tal y como se requiere para la planificación de red.

La manera más sencilla para estimar el tráfico para una aplicación comercial es con los minutos de uso mensuales (MoU, Monthly Minutes Of Use). El servicio se vende después a los usuarios en términos de cuotas por MoU, las cuales varían en función del tipo de usuario, tal y como se aprecia en la tabla 6.1:

Los requisitos de la planificación de red deben expresarse para la hora de mayor carga de tráfico, puesto que será el factor limitante en la red. Para poder llegar a una solución de compromiso entre cobertura y capacidad, los requerimientos deben desglosarse en diferentes áreas, en lugar de considerar la red completa.

La distribución geográfica de los usuarios puede estimarse con estudios demográficos. Además, la información se torna más accesible, ya que aparecen bases de datos públicas y comerciales ordenadas por códigos postales o categorizadas en función de su posición socio-económica.

Dado que la distribución de población puede estimarse a partir de estas fuentes, el resto del trabajo consiste en traducir los datos de MoU en horas de máxima ocupación, basándose en una estimación de la penetración del servicio.

Los requisitos de máxima ocupación se pueden estimar a partir de los datos de MoU calculando los factores de cresta entre días (interdía) y dentro de un mismo día (intradía). El factor de cresta entre días es la diferencia de carga de tráfico entre el día de máxima ocupación y el día medio (utilizando nuevamente datos estadísticos). Éste tiene un rango que oscila entre 1,2 y 3. Con esta información, el tráfico por usuario puede calcularse utilizando la fórmula 6.1.

$$Tráfico_{máx} = \frac{Tráfico_{mensual}}{30 \times 24} \times F_{cresta_{entredía}} \times F_{cresta_{intradía}} \quad (6.1)$$

Con la adopción de tarifas planas, en las cuales se ofertan gran cantidad de minutos a un coste fijo, además de descuentos en horario nocturno o fines de semana, los factores de cresta interdía o intradía se han reducido considerablemente durante los últimos años.

Una forma alternativa de estimar el tráfico de máxima ocupación a partir de MoU es asumir que el tráfico se distribuye según una distribución de Poisson. Este método puede utilizarse para estimar una distribución de tráfico irregular, incluso si no se conoce el factor de cresta. Utilizando la forma simplificada para la distribución de Poisson, la carga en horario de máxima ocupación se puede utilizar hallando el valor de «n» que satisface la ecuación 6.2.

$$\sum_{i=0}^n \frac{e^{-\lambda} \lambda^i}{i!} \leq (1 - GoS) \quad (6.2)$$

En dicha fórmula, λ representa la carga media, y «GoS» representa el grado de servicio deseado para ese sistema, o la probabilidad de que un determinado usuario experimente bloqueo. En este caso, «n» es coherente con las unidades de λ , bien MoU en horario de máxima ocupación, o miliErlangs (mErl) en horario de máxima ocupación.

MoU debe traducirse en mErl, el valor que resulta con frecuencia más inteligible para los planificadores de red. La conversión es sencilla: 1 Erlang corresponde a 1 recurso utilizado a lo largo de todo el periodo de observación. Durante el periodo de máxima ocupación, el periodo de observación es 1 hora, o 60 minutos. Por tanto, 1 Erlang = 60 MoU, o 1 MoU = 16,6 mErl.

La figura muestra un ejemplo numérico de esta distribución, en el cual se asume lo siguiente:

- ⇒ Todas las llamadas se atienden en un orden aleatorio.
- ⇒ Están disponibles un número infinito de recursos.
- ⇒ El tiempo de espera es constante o exponencial.
- ⇒ Todas las llamadas bloqueadas son eliminadas.

Los parámetros del ejemplo serán los siguientes:

- ⇒ 300 MoU por cliente.
- ⇒ Tráfico medio por hora, 0,41 MoU.
- ⇒ Escala temporal, 30 días por mes.
- ⇒ Observación diaria, 24 horas.
- ⇒ Tráfico medio equivalente por cliente, 6 mErl.

La ecuación se resuelve para $n=12$, significando que el tráfico en la hora de máxima ocupación se asume de 12 mErl para la planificación de red. Esto corresponde con un factor de cresta de 1,7. Este ejemplo asume que el tráfico está distribuido regularmente a lo largo de todos los días, aunque es más preciso suponer que el tráfico se distribuye irregularmente durante esos días de la semana. En el caso de una distribución de Poisson, puede seguirse un procedimiento de dos fases para determinar el tráfico esperado por día y el tráfico esperado en la hora de máxima ocupación. En base a los cálculos anteriormente realizados, el tráfico por cliente durante la hora de máxima ocupación es de 19 mErl. Esto está basado en un factor de cresta día a día de 1,6 y en un factor de cresta intradía de 1,6.

Otra manera de determinar los requerimientos de tráfico es utilizar la información existente en la red, en cuyo caso se conocen los requisitos de tráfico por cada celda, incluyendo distribución geográfica. Para tráfico CS (voz), extrapolar el tráfico WCDMA de otra red será válido siempre que

los cargos por servicio sean consistentes. Para tráfico de paquetes PS, la extrapolación es menos válida porque el tráfico tiende a depender del servicio ofrecido (máxima tasa de datos).

Mientras que el GoS depende de unas metas específicas determinadas por los operadores, los valores de GoS definidos para servicios de voz son normalmente en torno al 2% para usuarios de alta movilidad y en torno al 1% para usuarios de baja movilidad (servicios WLL, Wireless Local Loop Services). No debe confundirse este GoS con la probabilidad de éxito de una llamada; el GoS sólo tiene en cuenta los bloqueos (llamadas para las cuales no hay recursos disponibles), mientras que la probabilidad de éxito incluye cualquier tipo de fallo.

6.2. Descripción de los recursos

En cualquier caso, e independientemente del servicio, siempre que un usuario intente utilizar uno se deben desencadenar una serie de sucesos en los que se verán implicados casi todos los elementos de la red y que tendrán como objetivo, en primer lugar, dictaminar si es posible iniciar el servicio, en segundo lugar, reservar y mantener los recursos necesarios de la conexión y, por último, liberarlos al finalizar el mismo.

La decisión de si un servicio puede ser soportado en un determinado momento por la red dependerá de los recursos que consume cada uno, ya que no consume lo mismo una conversación de voz que una descarga de datos. Cabe entonces preguntarse de qué recursos se dispone en una red UMTS y en qué medida son empleados por cada tipo de servicio.

Los recursos objeto de estudio en esta Tesis, y por tanto susceptibles de estudiar cómo deben ser dimensionados, son los siguientes:

⇒ Hardware de los equipos: la modularidad de los equipos que se instalan en los emplazamientos provoca que no todos los nodos de la red tengan las mismas prestaciones. En esta Tesis se estudiará cómo intervienen en el dimensionamiento los recursos siguientes:

- Capacidad de procesado en banda base («Channel Elements»).
- Códigos de canalización («Spreading Codes»). Utilizados para aplicar las técnicas de espectro ensanchado propias de UMTS.

⇒ Capacidad del canal Iub.

En apartados posteriores se detallan más ampliamente sus características.

6.3. Empleo de los recursos

La característica más destacada de la red estudiada a lo largo de esta Tesis es que soporta diferentes servicios. Cada tipo de tráfico tiene sus propias características y su tratamiento está además definido de una manera diferenciada. No obstante, todas las peticiones de servicio compiten por los mismos recursos de la red, con lo que los algoritmos de control de admisión de la RNC pueden modelarse según el algoritmo de Kaufman-Roberts para redes multiservicio, que se estudiará detalladamente a lo largo de posteriores capítulos.

En este apartado se explica cómo los diferentes tipos de tráfico hacen uso de los recursos anteriormente comentados.

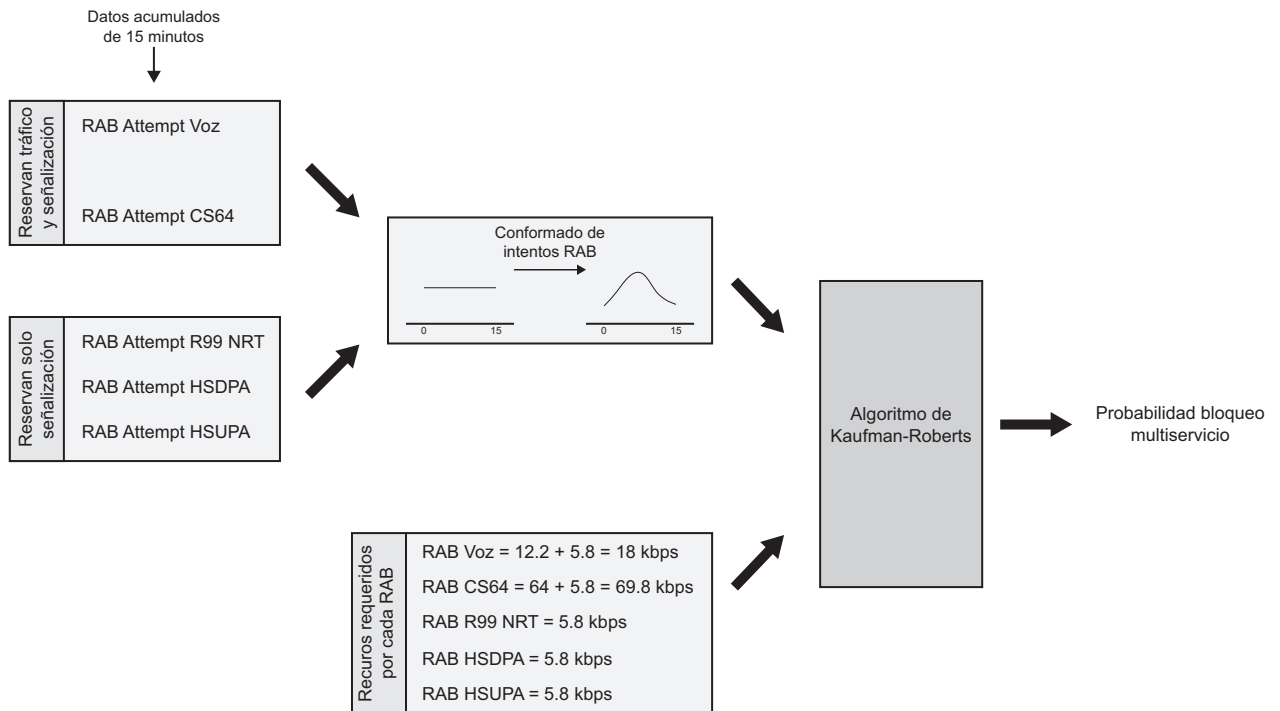


Figura 6.3: Esquema de la gestión de tráfico rígido

6.3.1. Tráficos rígidos

Las peticiones RAB llegan a la RNC. Cada servicio querrá reservar una conexión CBR diferente (R99 NRT, HSDPA y HSUPA reservarán recursos de señalización). A partir de ahí, el CAC empleará un algoritmo propietario para determinar si la conexión es admitida o denegada.

Dado que se trata de un escenario donde varios servicios demandan diferentes recursos, puede modelarse el CAC mediante el algoritmo de Kaufman-Roberts visto en el capítulo 4. Este algoritmo determina la accesibilidad del sistema en base a una serie de parámetros de entrada:

- ⇒ Uno de los parámetros de entrada requeridos para este algoritmo es el tráfico asociado a cada uno de los servicios, y será necesario modelarlo adecuadamente para imitar en la medida de lo posible el comportamiento real. Este aspecto será objeto de estudio en capítulos posteriores.
- ⇒ Los recursos que consume cada servicio.
- ⇒ La capacidad total disponible.

En el CAC siempre se supondrá que el tráfico a reservar es el máximo posible que puede soportar un servicio; si, por ejemplo, se tiene una solicitud RAB de voz, se buscará reservar los máximos recursos que consumirá, es decir, 12.2kbps más la señalización asociada de 5.8kbps. Sin embargo, el tráfico real medio será menor ya que no siempre se transmite al máximo posible. En el caso de la voz, cuando no se está hablando, se transmiten descriptores de silencio cuyo *bitrate* es inferior a 1kbps.

6.3.2. Tráficos elásticos

Dentro de los tráfico elásticos existe el R99 NRT, esto es, paquetes en formato R99, y tráfico HSPA. Por configuración de la red los tráfico se tratan como UBR+ y UBR respectivamente, por

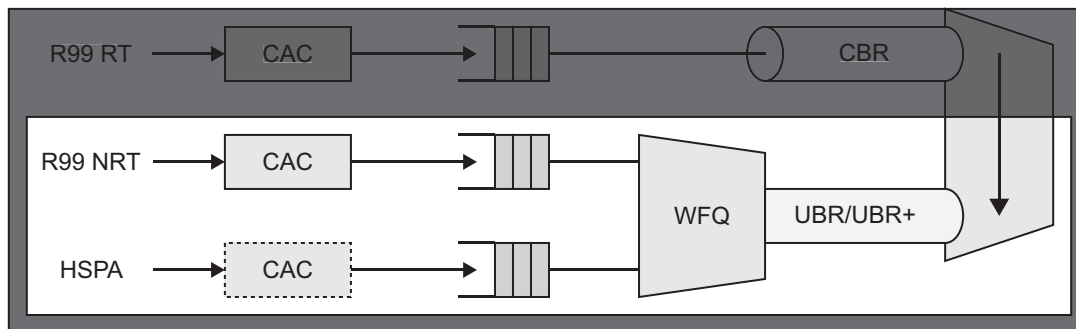


Figura 6.4: Esquema de la gestión de tráfico elástico

tanto, R99 NRT es prioritario sobre HSPA pero en ninguno de los dos casos existe prioridad sobre el tráfico R99 RT y, además, no se reserva ningún tipo de recurso.

A pesar de ser tráfico elástico, que en teoría nunca se bloqueará, se necesita un CAC dado que la señalización asociada sí que es un servicio rígido y que necesita de reserva de recursos. En la figura 6.4 se marca el CAC de R99 NRT con línea continua mientras que para HSPA se marca con línea punteada. Esto es así porque el CAC para R99 NRT es mucho más estricto: dado que las conexiones de este tipo tienen RABs de ancho de banda fijo (64kbps, 128kbps o 384kbps) se puede llegar al límite y tener que bloquear una conexión si la línea no tiene ancho de banda para alojar una nueva conexión de 64kbps. Sin embargo, esto último no sucede en HSPA, ya que el ancho de banda se reparte equitativamente entre todos los usuarios, así, el CAC solo necesita garantizar el flujo de señalización asociado.

En cuanto a la capacidad total disponible, en este caso se trabaja sobre ATM y, por tanto, la capacidad total disponible viene dado por el VCC destinado al transporte de tráfico rígido.

6.3.2.1. Consideraciones sobre R99 NRT

El tráfico de paquetes R99 NRT es un tanto especial: las redes Ericsson se pueden configurar para tratar a este tráfico bien como un servicio rígido, bien como un servicio elástico. La red bajo estudio, y siguiendo las últimas recomendaciones del fabricante, gestiona el tráfico R99 NRT como elástico. Esto quiere decir que se podrán producir situaciones donde el servicio de paquetes R99 vaya modificando el *bitrate* del RAB: así podría subirse de un RAB de 128 kbps a uno de 384kbps o bajar a otro de 64kbps.

Cuando se solicita un RAB lo que normalmente sucede es que se siga el algoritmo mostrado en la figura 6.5. En un primer momento, el CAC decide si el RAB tiene recursos suficientes en la red para ser cursado. Si es así, se comienza con un *bitrate* de 64kbps e inmediatamente después se intentará subir hasta el máximo de 384 kbps. Existe una limitación por códigos de un máximo de 8 usuarios a este *bitrate*. En caso de sobrepasarse, el CAC evaluará si existe capacidad suficiente para cursar el tráfico sobre 128 kbps. Si tampoco es posible, el RAB no se moverá y se cursará a la velocidad inicial de 64kbps.

Existe un punto de total incertidumbre en el esquema: es habitual que los operadores limiten el ancho de banda de los usuarios en virtud de su consumo y de las tarifas que tengan contratadas. Si, por ejemplo, la limitación es de 128 kbps, un usuario que hubiera sobrepasado el límite no podría optar a un RAB de 384 kbps, aun cuando no se hubiese llegado a la cifra de 8 conexiones de este tipo.

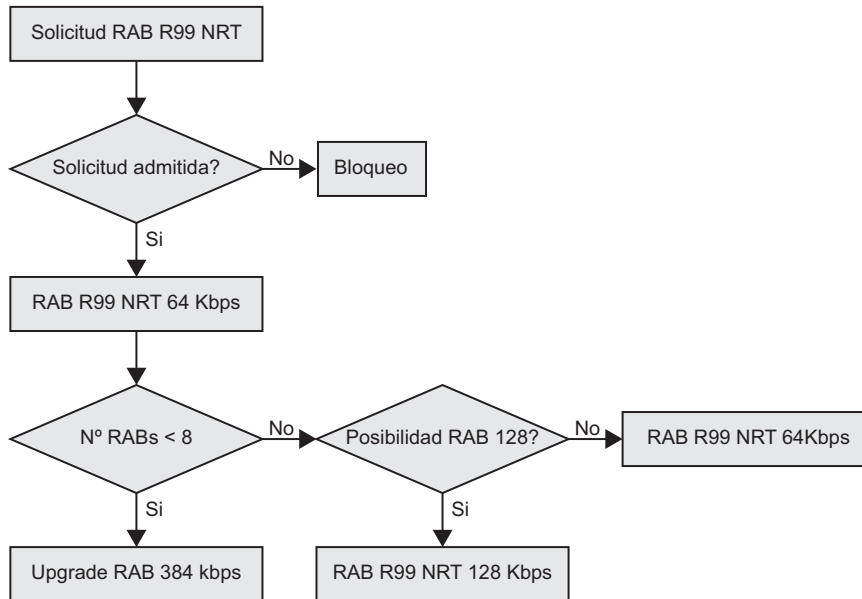


Figura 6.5: Esquema de la gestión de la solicitud de RAB R99 NRT

Como el tratamiento del tráfico R99 NRT es elástico, a medida que la carga de tráfico aumenta la RNC puede decidir bajar el bitrate de una conexión. En el caso de estar en máxima congestión es de esperar que todos los RABs R99 NRT estén siendo cursados a un bitrate de 64 kbps. Por tanto, y como lo que más interesa son las situaciones de congestión, siempre que se realice cualquier cálculo se considerará que todas las conexiones son a 64kbps, hecho que no será real cuando el canal Iub este vacío y libre congestión pero que será correcto durante los momentos de más tráfico.

El tráfico de paquetes R99 puede variar su ancho de banda entre 64kbps, 128kbps o 384kbps. La red Ericsson sobre la que se trabaja permite tan sólo un máximo de 8 conexiones a 384kbps. A partir de ese momento todas las conexiones que accedan al sistema lo deberán realizar sobre RABs de *bitrates* inferiores.

6.3.2.2. Control de Admisión en R99 NRT

En el punto 6.3.2.1 se vio que en R99 NRT existía un CAC en el que los RABs siempre partían de 64 kbps, a pesar de ser un servicio elástico que en teoría nunca se bloquea y no debería existir problema en poderse cursar. Sin embargo, un nodo B con un línea de transmisión formado por un solo E1, es decir, 2048kbps y considerando que solo cursa tráfico R99 y que la señalización es despreciable (no es real), tendría un máximo de conexiones a 64 kbps (ecuación 6.3).

$$\text{Nº RAB R99 NRT 64kbps máximo} = \frac{2048 \text{ kbps}}{64 \text{ kbps}} = 32 \quad (6.3)$$

El tráfico R99 NRT tiene prioridad sobre el tráfico HSPA: el primero se cursa sobre UBR+ y el segundo sobre UBR, es decir, que cuando se aplique el CAC el tráfico HSPA es como si no existiera. El único tráfico que existiría en el Iub sería el tráfico rígido. Si ocurriese el caso que una nueva conexión R99 NRT no tuviera sitio debido al tráfico HSPA, sería éste el que debería reducir su ancho de banda para dejar sitio al nuevo RAB R99. En la figura 6.6 se muestra el esquema del CAC para este tipo de RABs.

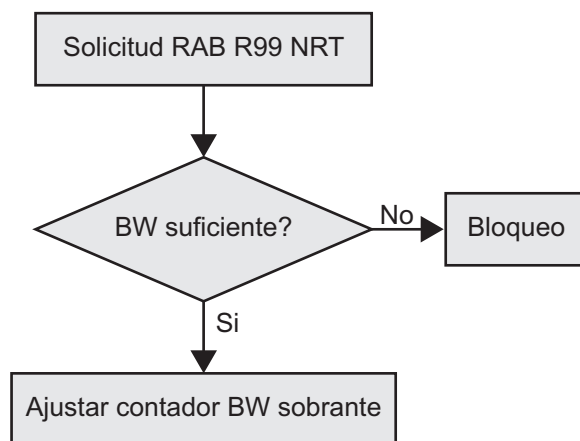


Figura 6.6: Esquema del funcionamiento del CAC de paquetes R99

6.3.2.3. Cola WFQ

En la transmisión de información entre RNC y Nodo B se emplea «Espera Equitativa Ponderada (*Weighted Fair Queuing* -WFQ). Se trata de una técnica de encolamiento que proporciona QoS en redes convergentes. Trata de evitar la congestión y controla directamente las colas de los nodos mediante un tratamiento diferencial del tráfico proporcionado por una determinada disciplina de servicio.

A la hora de realizar el modelo no se tendrá en cuenta, ya que en ningún caso se aborda el orden de transmisión de las tramas.

6.3.3. Tráfico HSPA

Tanto en los tráficos de voz como de videollamada y de paquetes R99, una vez que la conexión ha sido aceptada se mantienen las tasas de velocidad (*throughput*) fijas durante el tiempo que dure el servicio. Sin embargo, puesto que en HSPA esto no ocurre, y es precisamente el *throughput* el objetivo para este servicio, debe poder modelarse de alguna manera su comportamiento.

HSPA aprovecha el tráfico sobrante del resto de conexiones y se conoce el número de conexiones activas y su duración, ya que ha sido calculado en el punto anterior (CAC). Sin embargo, no se ha calculado el verdadero *throughput* que consumen ya que, por ejemplo, una conexión de voz empleando el codec AMR 12.2kbps no consume realmente 12.2kbps en el canal Iub: esto es debido a que dicho *codec* no transmite nada cuando el interlocutor permanece sin hablar (se transmite en su lugar un descriptor de silencio para ahorrar ancho de banda). Además toda conexión que viaje por la interfaz Uu se encapsula en celdas ATM que se transmiten por el Iub.

Así pues, es necesario calcular el tráfico real que circula, lo cual se expone detalladamente en el apéndice D.

6.4. Códigos de canalización

UMTS emplea la tecnología *Wideband Code Division Multiple Access* (WCDMA - Acceso múltiple por división de código de banda ancha). Esta técnica permite que, sobre una única portadora de 5 MHz común en toda la red, diferentes usuarios puedan intercambiar información con los diferentes nodos B de forma simultánea, todo ello gracias a que antes de transmitir cualquier secuencia

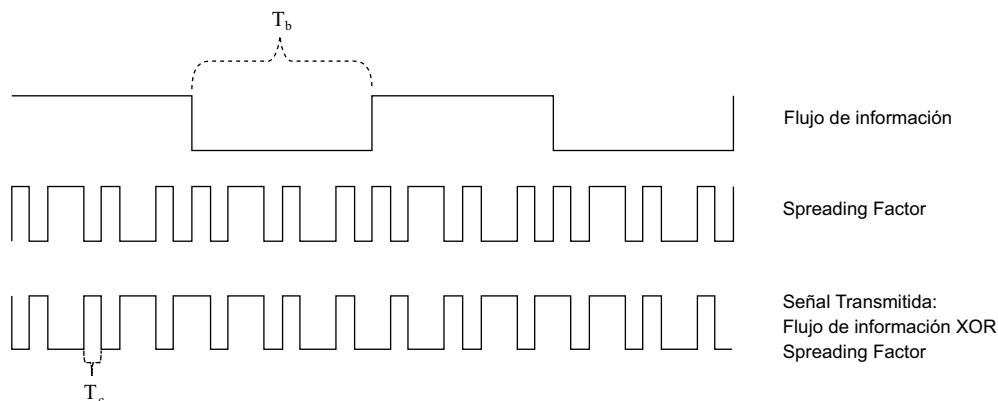


Figura 6.7: Ejemplo de ensanchamiento de la señal en WCDMA

de bits se produce un ensanchamiento del espectro mediante el uso de códigos de canalización, cuya característica más importante son los códigos de ensanchamiento o *spreading codes* (SC).

A modo de ejemplo, se muestra el ensanchamiento de la señal en la figura 6.7.

En WCDMA para cada enlace de un UE (*User Equipment*) con un Nodo B se reserva un código de *Scrambling* y un código de *Spreading*. El producto del código de *Scrambling* con el código de canalización debe ser único para conseguir una transmisión libre de interferencias. Es importante señalar que en el enlace ascendente los UE se distinguen de forma unívoca por su código de *Scrambling*, es por esto que cada usuario puede hacer uso de cualquier código de *Spreading* sin preocuparse por las interferencias en el enlace ascendente. Sin embargo en el enlace descendente los datos son transportados por el mismo nodo B a un número de usuarios. En este caso, el mismo código de *Scrambling* puede ser utilizado para transportar datos en diferentes canales. Por ello, para asegurar que el producto de los códigos sea único en el enlace descendente se necesita que cada código de canalización sea diferente a todos los demás. Es por esto que la gestión de los códigos de *Spreading* en el enlace descendente es crucial.

La figura 6.8 muestra el procesado de la señal tanto en el emisor como en el receptor. Los dos primeros pasos: producto con código de *Spreading* y con código de *Scrambling*, son las operaciones necesarias para ensanchar la señal. Finalmente se sube la misma a la frecuencia de portadora.

El ensanchamiento de la señal es una técnica de modulación empleada para la transmisión de datos digitales. Consiste en «ensanchar» la señal a transmitir a lo largo de una banda muy ancha de frecuencias, mucho más amplia que el ancho de banda mínimo requerido para transmitir la información que se quiere enviar.

Las ventajas del espectro ensanchado son varias:

- ⇒ Aumento importante de la resistencia a todas las interferencias.
- ⇒ Elimina las interferencias multitrayecto.
- ⇒ Confidencialidad de la información transmitida gracias a los códigos pseudoaleatorios.

Los códigos de canalización se generan a partir de un árbol de códigos.

Los códigos en diferentes niveles tienen SF (Factor de *Spreading*) diferentes, lo que posibilita la transmisión multi-tasa. El código es escogido de una capa de tal forma que el producto del SF y la tasa de datos sea constante para todas las clases de tráfico.

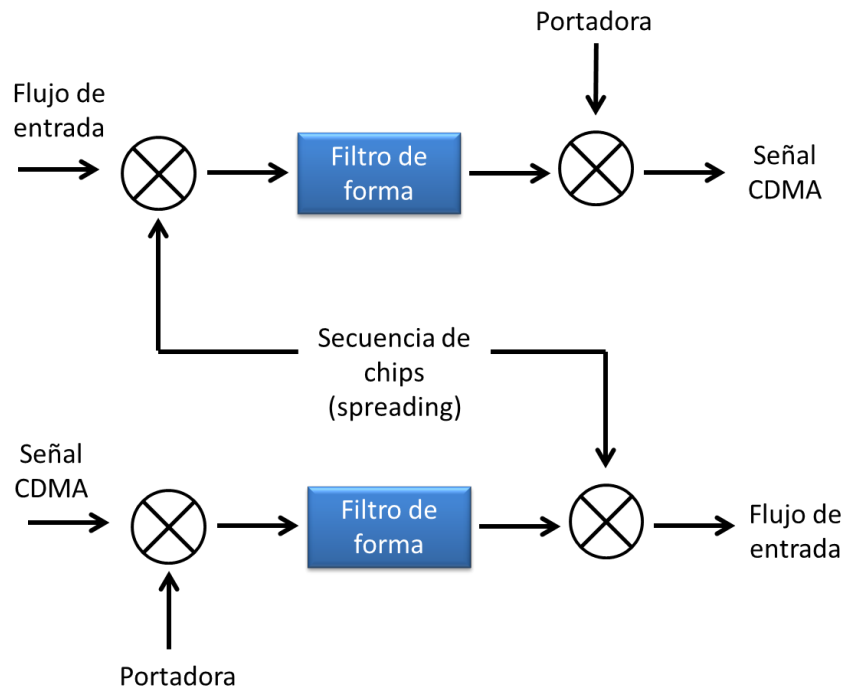


Figura 6.8: Proceso de ensanchamiento de la señal

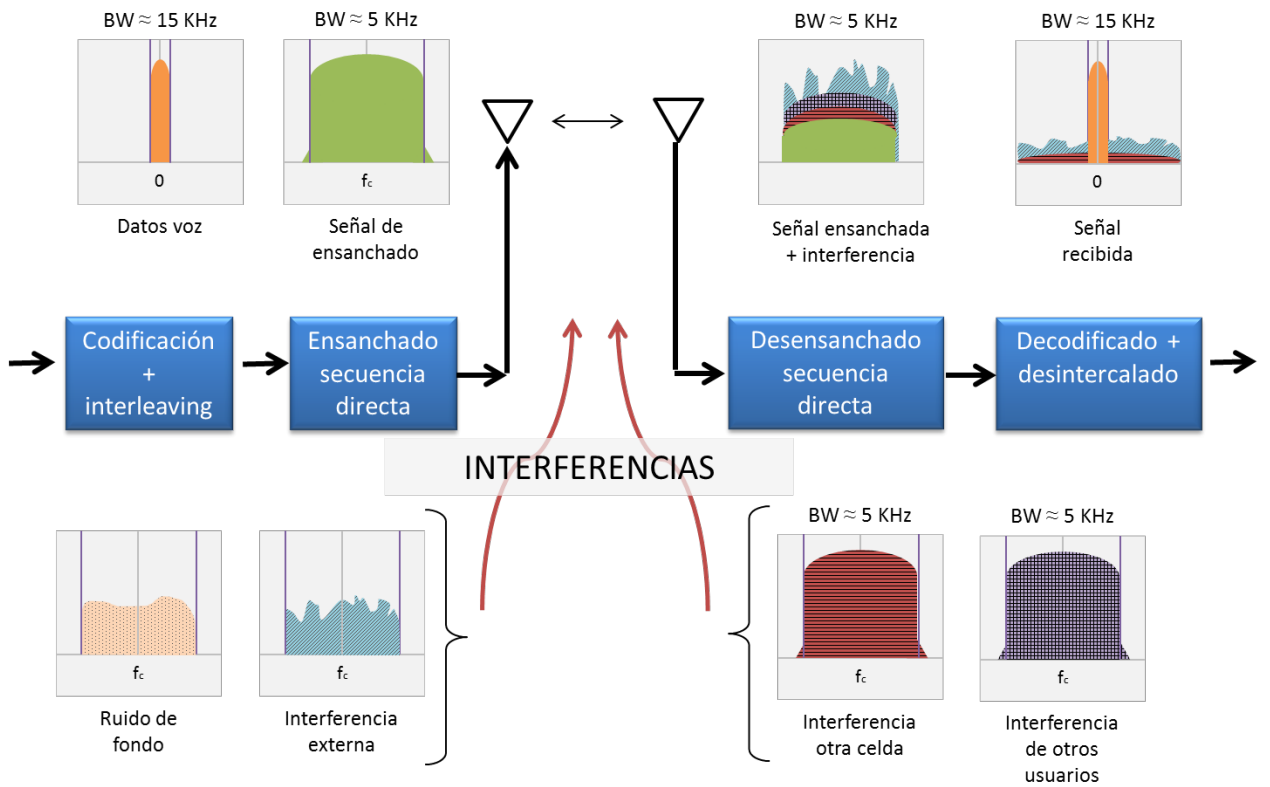


Figura 6.9: Ensanchamiento de la señal

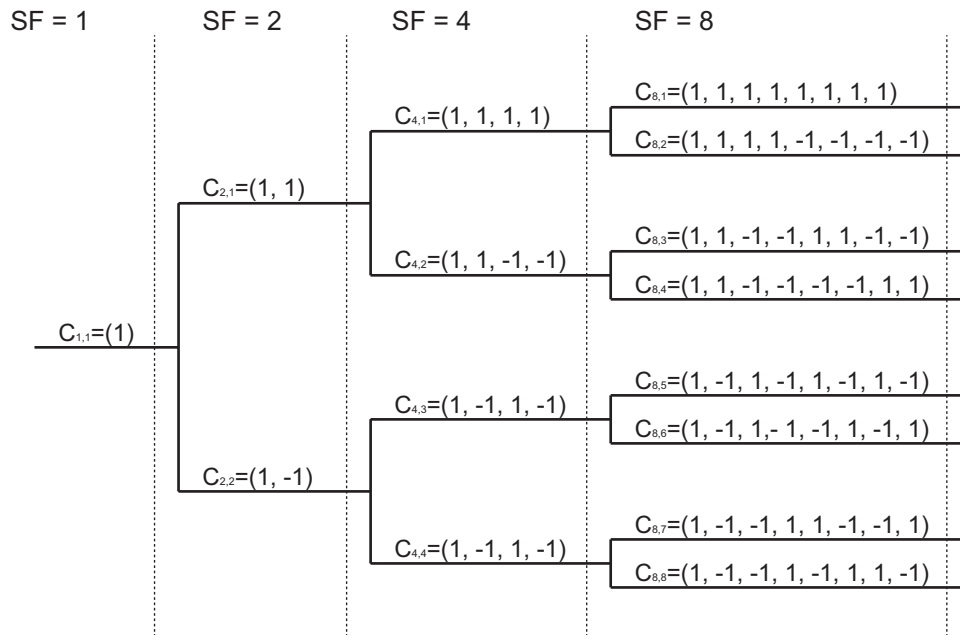


Figura 6.10: Ejemplo de árbol de códigos de ensanchamiento.

En el árbol de códigos, cada código de una capa es ortogonal a todos los demás en la misma capa. Sin embargo no es ortogonal ni con sus antecesores ni con sus descendientes. Al ocupar un código, tanto sus antecesores como sus descendientes quedan bloqueados.

El funcionamiento de este árbol de códigos puede entenderse mejor con el ejemplo de la figura 6.11. En ella puede verse un árbol de solo 4 niveles en el que están ocupados 2 códigos (en negro): uno de nivel 4 y otro de nivel 3. Por el simple hecho de ocupar estos 2 códigos estarán bloqueados los otros 6 códigos que vemos en gris, sus antecesores y descendientes.

El uso de códigos de Spreading y de Scrambling es diferente en el enlace ascendente y descendente. En el enlace descendente los códigos de Spreading son usados para multiplexar diferentes canales transmitidos en la misma celda. En el enlace ascendente se usan para separar los canales de control y datos de un usuario específico. Los códigos de Scrambling, usando secuencias pseudo-aleatorias, se usan también para ensanchar la señal. En el enlace descendente, los diferentes códigos de Scrambling separan diferentes celdas, en el ascendente separan diferentes usuarios.

Es importante destacar que los códigos de canalización no suponen un factor limitante en el enlace ascendente, ya que cada usuario genera su propio árbol de códigos, por lo que no se trata de un recurso compartido. Por tanto, a la hora de abordar el estudio de las accesibilidades deberá tenerse en cuenta que en el caso del enlace ascendente, la accesibilidad debida a este factor es del 100 %.

6.5. Channel Elements

Otro de los recursos de la red está relacionado con la capacidad de procesamiento de los nodos B, encargados de modular y codificar las señales procedentes de los terminales de usuario. Se trata de un recurso limitado porque evidentemente éste tiene una capacidad máxima de servicios que puede atender.

Se define el «Channel Element» (CE) como una unidad de medida de los recursos de procesamiento

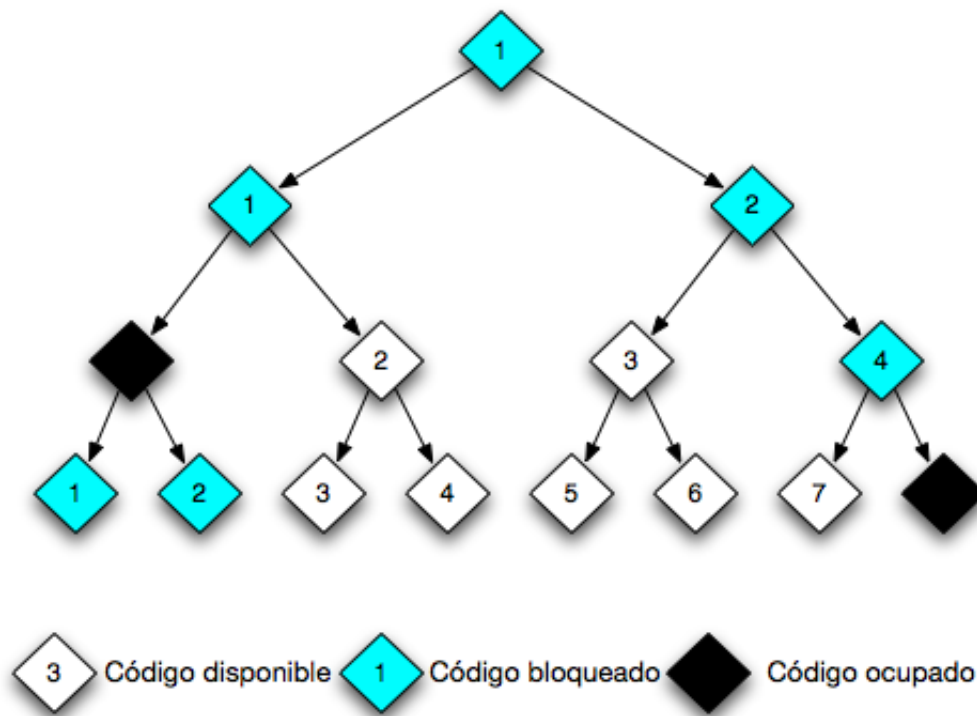


Figura 6.11: Árbol de códigos de cuatro capas

en banda base. En concreto, la definición normalizada es la siguiente:

- ⇒ 1 CE en *downlink* es igual al consumo de una conexión de voz empleando el codec AMR 12.2kbps con un árbol de códigos SF=128.
- ⇒ 1 CE en *uplink* es igual al consumo de una conexión de voz empleando el codec AMR 12.2kbps con un árbol de códigos SF=64.

Así pues, dependiendo de la capacidad de procesamiento de los equipos hardware tendremos un número máximo de conexiones máximas simultáneas y, por tanto, los CE son también un punto posible de congestión del sistema.

Hay 2 tarjetas dentro de los Nodos B que gestionan su capacidad:

- ⇒ Tarjeta DL: Control de la capacidad en el enlace descendente (*downlink*: del nodo al móvil).
- ⇒ Tarjeta UL: Control de la capacidad en el enlace ascendente (*uplink*: del móvil al nodo).

Los Channel Elements son las unidades de capacidad de estas tarjetas, de forma que se puede decir que una tarjeta DL tiene 128 CE o que una tarjeta UL tiene 256 CE. Cada servicio atendido consume un número diferente de CE en cada tarjeta, por ejemplo una llamada de voz consume 1 CE en la tarjeta DL y otro CE en la tarjeta UL, y un servicio de Circuitos Conmutados (por ejemplo: una videoconferencia) consume 4 CE en cada una de las tarjetas. En la figura 6.12 puede verse una representación del diferente consumo de recursos en un modelo de tarjeta.

Es importante destacar que se pueden instalar varias tarjetas tanto para el enlace ascendente como descendente, y para cada enlace hay diferentes modelos de tarjetas con mayor o menor número

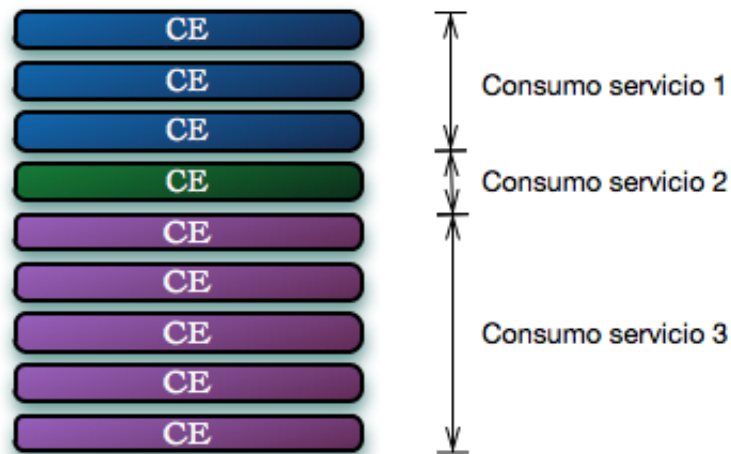


Figura 6.12: Consumo de recursos en tarjeta de nodo B

CE	Servicio
1	Voz
4	CS384
16	PS384

Cuadro 6.2: Consumo de CE en tarjeta DL

de CE. Los distintos servicios ofrecidos por la red 3G no consumen el mismo número de CE en los dos tipos de tarjetas, aunque es muy parecido. Las tablas 6.2 y 6.3 muestran el consumo de recursos que tienen los distintos tipos de servicios, tanto en el enlace ascendente como en el descendente.

Para el servicio de HSPA el consumo de CE funciona de manera diferente y más compleja. Para su gestión utiliza otro recurso de las tarjetas DL: los ResourceIds.

- ⇒ Una tarjeta DL tiene 3 ResourceIds.
- ⇒ Cada ResourceId puede ofrecer hasta 20 códigos HSDPA.
- ⇒ Un sector puede tener 15 códigos HSDPA como máximo.
- ⇒ Varios sectores pueden compartir el mismo ResourceId (códigos).

La gestión de los ResourceIds depende del modelo de tarjeta considerado. En particular, para el caso del fabricante Ericsson, el lector puede acudir al Apéndice F.

CE	Servicio
1	Voz
2	CS384
8	PS384

Cuadro 6.3: Consumo de CE en tarjeta UL

Denominación	Kbps
E1	2,048
E2	8,448
E3	34,368
E4	139,264

Cuadro 6.4: Multiplexación PDH en Europa

6.6. Iub sobre ATM

De todas las interfaces que existen en la arquitectura UMTS, posiblemente la más crítica desde el punto de vista del despliegue de red de un operador sea el canal Iub. Mediante éste se establece comunicación entre los nodos B (encargados de dotar de cobertura a la red) y la RNC.

El interés en el estudio del Iub viene dado por el importante impacto económico que supone su despliegue: en un país como España un operador puede contar con unas pocas decenas de RNCs y varios cientos de nodos B, por tanto, se necesita una red muy extensa y de gran capacidad para hacer concentrar toda la información de los nodos B dispersos en las RNC.

Sin un dimensionamiento correcto, podrían darse situaciones de congestión en el Iub que afectarían gravemente al funcionamiento de decenas de nodos, los usuarios percibirían que sus llamadas y conexiones de datos no se podrían realizar, el operador dejaría de facturar tráfico y lo que es aún peor, las quejas y bajas de usuarios aumentarían debido a la mala calidad del servicio.

Además, en una situación de congestión del Iub, las medidas a tomar son (además de costosas) dilatadas en el tiempo: el despliegue o ampliación de la red requiere de un diseño o rediseño de la misma, tiempos de pedido y suministro de equipos y solicitud de licencias y permisos de diversa índole. En el caso de la congestión por SF y CE, ésta sólo afectaría a un nodo B y podría ser solucionada mediante la ampliación de sus tarjetas *hardware* o, en el peor de los casos, con el despliegue de un nuevo nodo B vecino. Éste absorbe parte de la carga del congestionado, lo que sin duda es mucho más sencillo, menos costoso y más rápido que en el caso de congestión en el Iub.

El transporte de la información en el Iub hace uso de la tecnología PDH (Jerarquía Digital Pleisícrona). PDH se basa en canales de 64 kbps. En cada nivel de multiplexación se van aumentando el número de canales sobre el medio físico. Es por eso que las tramas de distintos niveles tienen estructuras y duraciones diferentes. Además de los canales de voz en cada trama viaja información de control que se añade en cada nivel de multiplexación, por lo que el número de canales transportados en niveles superiores es múltiplo del transportado en niveles inferiores.

Existen tres jerarquías PDH: la europea, la norteamericana y la japonesa. La europea usa la trama descrita en la norma G.732 de la UIT-T mientras que la norteamericana y la japonesa se basan en la trama descrita en G.733.

La tabla 6.4 muestra los distintos niveles de multiplexación PDH utilizados en Europa.

Del mismo modo que en el Nodo B la capacidad se mide en CE, en el Iub se mide en Kbps. Y del mismo modo que en el nodo B se aumentan las capacidades instalando tarjetas HSTX o RAXB que aporten más CE, en el Iub se aumenta la capacidad incrementando el número de E1 en el mismo.

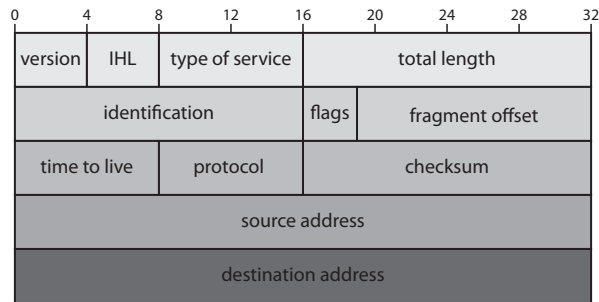


Figura 6.13: Cabeceras IPv4

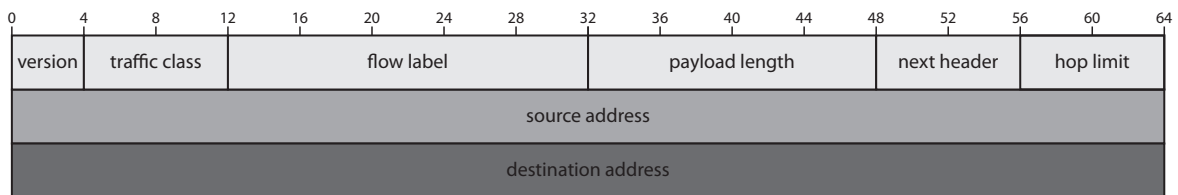


Figura 6.14: Cabeceras IPv6

6.7. Iub sobre IP

Recientemente se ha venido imponiendo el uso de la transmisión IP, principalmente sobre nodos B en zonas urbanas, y la tendencia en los próximos años parece ser que es a incrementar el número de los mismos.

6.7.1. Bloqueo en transmisión IP

Cuando se emplea tecnología IP se pierden las prioridades de tráfico que ofrecía ATM: no existen conexiones CBR, UBR+ o UBR. Si se emplea IPv4 (caso de la red Ericsson en estudio) la calidad de servicio se marcará de alguna manera utilizando el campo «*type of service*» de la cabecera IP (figura 6.13). En caso de utilizar IPv6 se pueden incluir mejores mecanismos de QoS mediante el uso de las opciones de extensión que incorpora en la cabecera (figura 6.14).

En cualquier caso, se utilice IPv4 o IPv6 el funcionamiento será similar. Toda la información, ya sea voz, videollamada, paquetes R99 o tráfico HSPA viajará encapsulada en tramas IP, cada una con un descriptor de prioridad en la cabecera IP. Cuando se quiera iniciar un servicio será necesario pasar por el CAC de la RNC, con la salvedad de que en este caso no existen VCCs para los tráficos rígidos. El esquema será similar a excepción de un detalle, no se conocerá la capacidad total del Iub y será necesaria una estimación, que en la mayoría de los casos se proporciona a la RNC vía software.

El desconocimiento de la capacidad en el Iub viene dado en gran medida por los medios de transmisión empleados sobre IP. Normalmente, la capa 2 que transporta IP es Ethernet. La transmisión se realiza mediante fibra óptica, radioenlaces o xDSL. Dependiendo de estos medios, existen «cuellos de botella» que hacen en la mayoría de los casos impredecible la capacidad real disponible en el Iub:

- ⇒ Fibra óptica: cuando se emplea transmisión por fibra óptica la capacidad de ésta es muy superior a la capacidad de *switching* de los propios equipos, por tanto, muchas veces el límite no lo impone el medio físico, sino estos últimos.

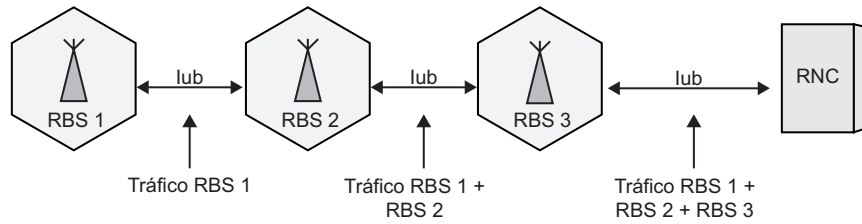


Figura 6.15: Agregado de tráfico

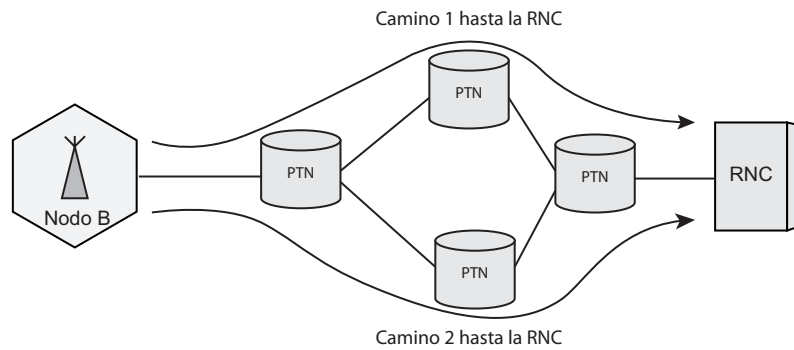


Figura 6.16: Enrutado de tráfico

⇒ xDSL: La tecnología xDSL es muy sensible a los fenómenos de diafonía que se producen en el par de cobre, por lo que la velocidad de sincronismo de estos equipos es variable.

⇒ Radioenlaces: el empleo de radioenlaces Ethernet va ligado siempre al empleo de modulaciones adaptativas que, como ya se muestra en la figura 3.18, poseen un ancho de banda variable con las condiciones radio.

Aun suponiendo que se pudiera determinar con precisión cuál es la capacidad real de la línea existe un problema mayor: la agregación de tráfico y el enrutado.

6.7.2. Agregación y enrutado de Tráfico

Al igual que sucede en Internet, las redes IP poseen la capacidad para buscar caminos alternativos para llegar desde un punto a otro. Además, y a diferencia de lo que sucedía en ATM donde los tráficos eran independientes y tenía sus propias «autopistas» con «carriles» independientes, aquí los tráficos comparten todos el mismo canal, de tal manera que ni se conoce exactamente el camino que van a seguir siempre los paquetes (figura 6.16) ni se sabe cuál es la capacidad sobrante que hay en el canal si no se conoce el tráfico que tienen todos los nodos que cuelgan de la RNC, ya que el tráfico se va agregando, tal y como se muestra en la figura 6.15.

6.7.3. Sentido del Bloqueo en IP

Visto los problemas que ocasiona el agregado y el enrutado del tráfico, cabe preguntarse por el sentido del modelado del canal Iub en este caso. El bloqueo que se pueda ocasionar debido al canal Iub dependerá no sólo del tráfico al que se somete el nodo B, sino que vendrá dado por el tráfico de una determinada área geográfica en la que varios nodos comparten parte de la transmisión y, por tanto, del Iub.

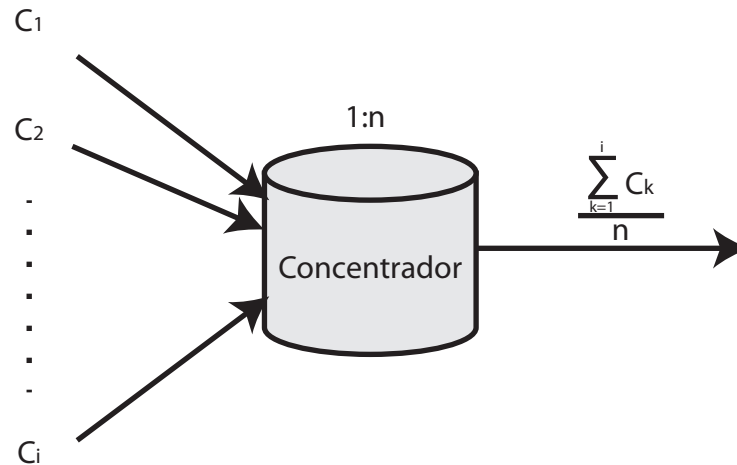


Figura 6.17: Esquema del dimensionamiento de un salto de red en IP

Podrían existir casos en los que un Nodo B presenta congestión en el Iub con bajo tráfico debido a un aumento del mismo en un nodo que comparta parte de la transmisión. Esto es así porque en situaciones de congestión del Iub lo que sucederá es que se irán descartando aquellos paquetes con menor prioridad en los diferentes puntos de conexión de la red (Radioenlaces, SMX y PTNs).

6.7.4. Dimensionamiento del Iub en IP

Debido a la agregación y enrutado de tráfico, la complejidad a la hora de dimensionar el Iub aumenta y harían falta modelos complejos donde se considerase el tráfico de áreas geográficas amplias, la capacidad de los equipos, las rutas posibles, etc.

Cuando se habla de emplear IP es porque normalmente se dispone de gran capacidad en el Iub, lo que, en resumidas cuentas, implica que el tráfico de datos ocupará un porcentaje muchísimo mayor al que ocupa voz, videollamada y señalización (la llamada banda ancha móvil). De hecho, el dimensionamiento no se hará para conseguir una determinada probabilidad de bloqueo, sino para lograr un GoS concreto en cuanto a caudal de tráfico por usuario se refiere. Actualmente, se está comenzando a desplegar nodos con celdas de capacidad 42 y 84 Mbps empleando HSPA+, lo que para nodos trisectoriales implicaría estar hablando de 126 y 252 Mbps. La capacidad de transmisión de estos nodos debe ser, pues, muy elevada.

Si se une la dificultad del diseño con el aumento constante de la demanda de tráfico por parte de los usuarios y el avance que se está produciendo en la capacidad de la interfaz aire Uu, resulta que cada poco tiempo es necesario ampliar la red para satisfacer las necesidades crecientes. Este es el motivo por el cual el dimensionamiento del Iub en IP se realice mediante un simple sistema de relaciones de la capacidad entrada/salida a un nodo.

Se empieza el dimensionamiento desde los elementos más externos de la red hasta la RNC, se define el GoS que se desea tener en estos nodos, por ejemplo, 30 Mbps. A partir de ahí, a cada nodo concentrador de tráfico se le aplica la siguiente regla: la suma de la capacidad total entrante es igual a la capacidad total de salida dividida entre un coeficiente. Este coeficiente o relación entrada/salida al nodo es un valor empírico, típicamente relaciones 1:2, 1:3, 1:4 o 1:5. La figura 6.17 muestra esta idea. A partir del primer nodo, se va escalando hasta llegar a la RNC.

Sirva a modo de ejemplo el dimensionamiento hecho en la red mostrada en la figura 6.18.

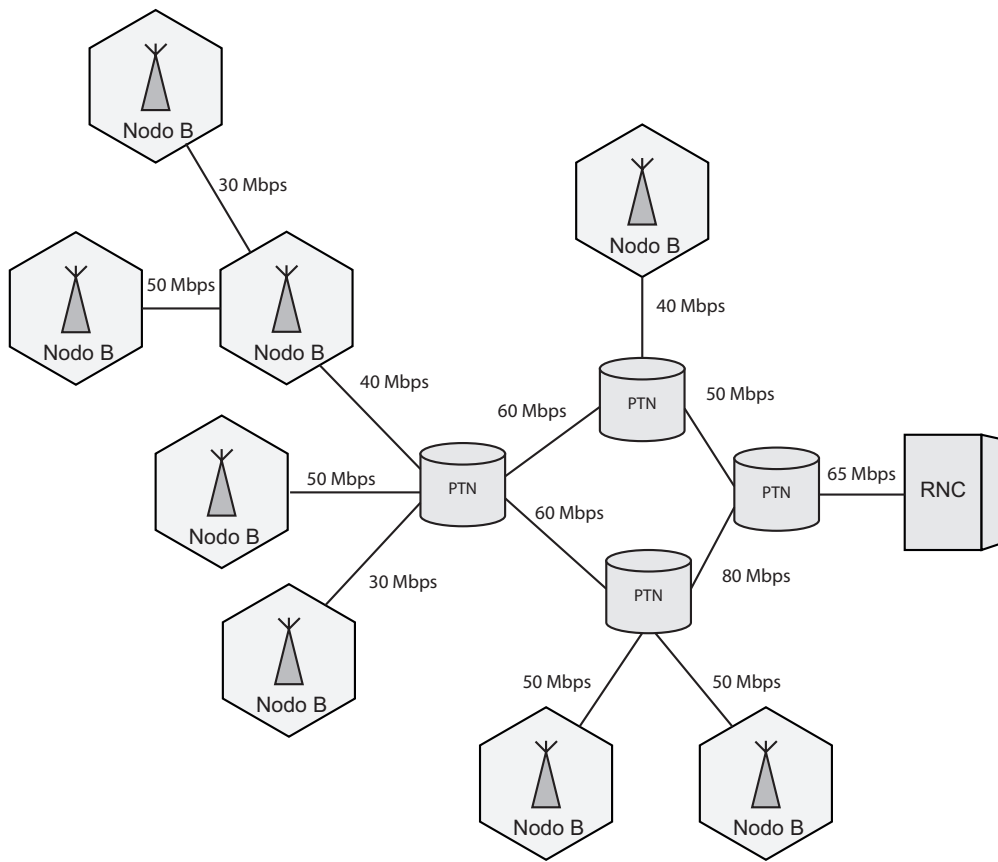


Figura 6.18: Ejemplo de dimensionamiento IP

Normalmente, el dimensionamiento se realiza en el caso peor, que será si la transmisión es mediante radioenlaces Ethernet en el momento en el que estén trabajando con la modulación más baja programada.

6.8. Estudio de las interferencias en el Uplink

6.8.1. Capacidad del Uplink y parámetro N_{pole}

El demodulador que se encuentra en el nodo B sólo puede realizar su trabajo con éxito cuando la relación señal a ruido es superior a un determinado valor umbral, de manera que si por alguna razón aumenta la intensidad de ruido, no será posible la demodulación. Este incremento de ruido puede venir ocasionado, entre otras causas, por las interferencias causadas por las celdas vecinas, pero también muy importante, por tratar de albergar un gran número de usuarios.

Así, los usuarios que se encuentran más alejados de la estación son rechazados, o lo que es lo mismo, el radio de «cobertura» de la celda se ve limitado. Por tanto, si se incrementa progresivamente el número de usuarios, aumentará el nivel de ruido y por consiguiente la cobertura tiende a cero.

El ruido en el Uplink se incrementa con la carga de manera exponencial. Cuando este ruido tiende a infinito, entonces no pueden añadirse más usuarios a la celda: la carga de la misma será del 100 % y se dice que ha alcanzado su «capacidad asintótica» (*Pole Capacity*). En las ecuaciones que siguen en este capítulo, se define el nº de usuarios máximo que soporta en estas condiciones con la notación N_{pole} .

Matemáticamente, para calcular la capacidad asintótica del enlace ascendente puede utilizarse la fórmula 6.4:

$$N_{pole} = \frac{W/R_b}{(1+i) \times \nu \times 10^{(E_b/N_0)/10}} \quad (6.4)$$

La definición de cada una de las variables que aparecen en la ecuación pueden encontrarse en la tabla 6.5.

- ⇒ W (*Spreading Bandwidth*). El ancho de banda es fijo y tiene un valor estándar de 3,84 megachips por segundo.
- ⇒ R_b (*Radio Bearer bitrate*). La capacidad puede incrementarse sólomente utilizando aplicaciones que requieran un bitrate menor. Un ejemplo, para la voz, es la utilización de AMR (*Adaptative Multirate*) con un valor de 7,95 frente al AMR clásico de 12,2. Dado que todos estos cambios degradarán en mayor o menor medida la calidad percibida por el usuario, deben probarse previamente.
- ⇒ E_b/N_0 (Relación energía por bit de señal y ruido total). Esta relación se ve influenciada enormemente por la tasa de datos, las condiciones del canal, la codificación de éste, cuál será el objetivo de BLER (*Block Error Rate*) y de la implementación *hardware*. En la tabla 6.6 se muestran algunos valores que pueden ser objetivo de diseño en el momento de la planificación.
- ⇒ ν (factor de actividad). Este factor depende de varios factores: la codificación del canal, las aplicaciones que tenga o, en el caso de la voz, del tipo de vocoder que se utilice. Para voz, se considera que un monólogo tiene un factor de actividad del 85 %, mientras que para el caso de una conversación normal en ambos sentidos, la actividad de la voz por usuario se estima en

torno al 40-45 % en la capa de aplicación. Para trasladar este resultado al factor de actividad debe tenerse en cuenta el esquema de codificación del canal y el vocoder utilizado. El vocoder es importante porque también codifica los silencios y el ruido de fondo, aunque a una tasa menor. El esquema de codificación también es importante porque el canal de transporte sólo puede manejar un número finito de formatos, para los cuales se añaden una portadora de señalización. Una vez factorizados, se estima que para una conversación normal, en ambos sentidos, el factor de actividad de voz se encuentra en torno al 58-67 %, con un posible pico del 85 % si no se utilizasen canceladores de eco. Considerando este rango, la actividad de voz de móvil a móvil requiere un factor de actividad mayor que el tráfico de móvil a «tierra». Para aplicaciones de datos PS, se puede definir también un factor de actividad (también conocido como factor de eficiencia) y dependerá del tipo de aplicación (e-mail, navegación web, streaming...) y del mecanismo de conmutación de canal que se haya implementado.

⇒ i (factor de interferencia). Depende principalmente de la calidad de la planificación de la red, porque representa las interferencias con celdas vecinas. El factor de interferencia está vinculado al solapamiento entre celdas y a la capacidad de una celda de controlar la potencia de una llamada. Este último punto se hace significativo cuando los tamaños del «active-set» están limitados artificialmente, o cuando se retrasa el handover. Los valores de i se encuentran en el rango de 0.55 a 0.65. Este rango representa una red optimizada con el mínimo solape entre celdas. Desafortunadamente esta optimización no es siempre compatible con las altas pérdidas por penetración en edificaciones que se esperan de las redes actuales. En las herramientas de planificación comerciales, este factor i no está disponible de forma directa, normalmente se tienen el factor de reutilización de frecuencias o la eficiencia de reutilización de frecuencias. Estas cantidades se pueden estimar a partir de la potencia total recibida y de la contribución a esa potencia de la celda misma, tal y como se representa en las ecuaciones 6.5 (factor de reuso de frecuencia) y 6.6 (factor de eficiencia de reuso).

$$F = 1 + i = \frac{P_{Rx,total}}{P_{Rx,celda}} \quad (6.5)$$

$$F_f = \frac{1}{1 + i} = \frac{P_{Rx,celda}}{P_{Rx,total}} \quad (6.6)$$

La figura 6.19 ilustra el concepto que se encuentra detrás de la ecuación 6.4. Cada usuario transmite a la máxima potencia para vencer el ruido, al mismo tiempo que la cobertura en el uplink va tendiendo a cero. A partir de la capacidad asintótica, se puede calcular una capacidad práctica N_{user} después de que se determine la carga en el uplink, tal y como se representa en la ecuación 6.7.

$$N_{usr} = N_{pole} \times \eta \quad (6.7)$$

La carga máxima en el uplink se determina de tal manera que asegure que la red se encuentra estable y que la cobertura no se ve afectada de manera adversa.

La capacidad calculada con las ecuaciones 6.4 - 6.7 representa el número de recursos disponibles en el Uplink. En la figura 6.20 se muestra gráficamente cómo a medida que crece el n^o de usuarios conectados se incrementa la interferencia en la celda, si bien a partir de un determinado valor N_{pole} el crecimiento de la interferencia es exponencial y provoca la sobrecarga en el nodo.

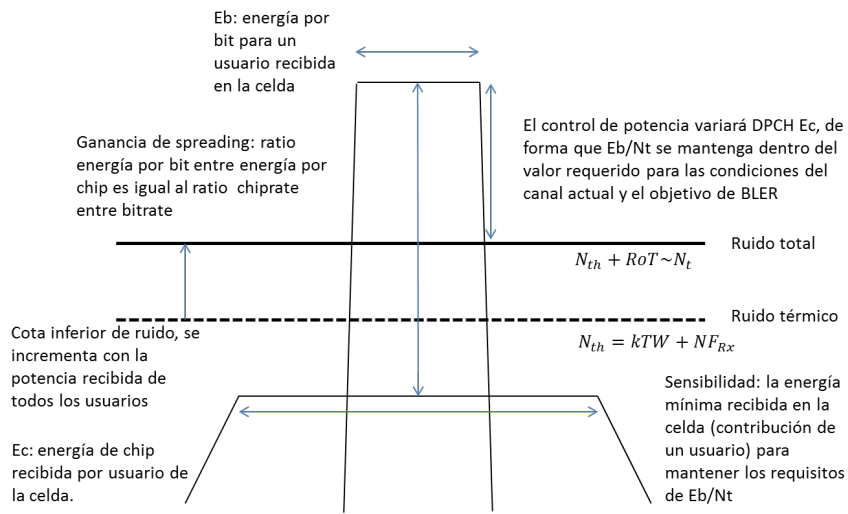


Figura 6.19: Conceptos de carga y capacidad en el Uplink

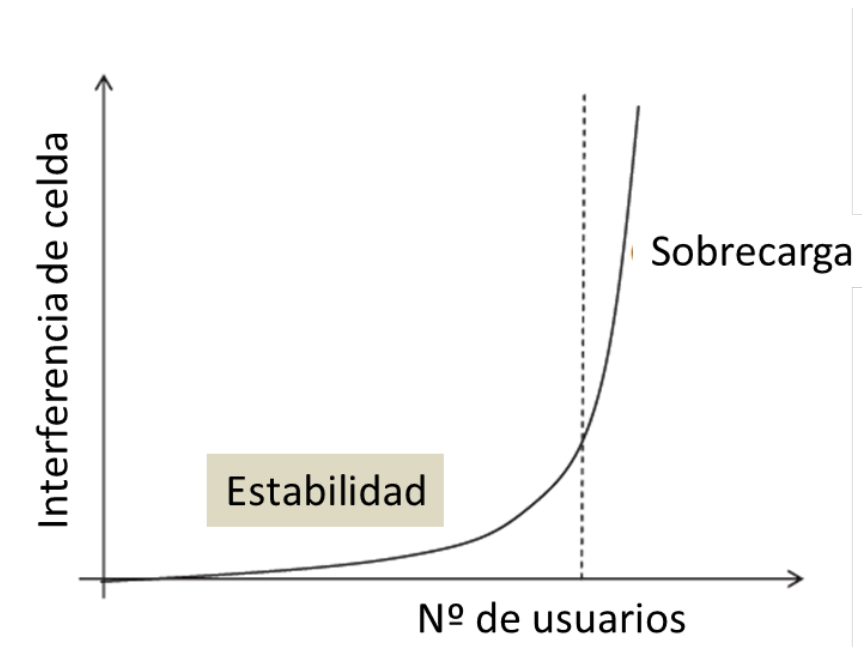


Figura 6.20: Incremento de la interferencia de celda con el nº de usuarios y sobrecarga

6.8.2. Factor de carga en el Uplink

La eficiencia espectral teórica de WCDMA puede calcularse a partir de la ecuación de carga que se presentará a continuación. Se define en primer lugar el parámetro E_b/N_0 , energía por bit dividido entre la densidad espectral de ruido (ecuación 6.8), que depende de la ganancia de procesamiento, la señal del usuario «j» y la potencia total recibida excluyendo la propia señal.

$$(E_b/N_0)_j = G_{procesado,usuario_j} \times \frac{Señal_j}{P_{t,Rx}} \quad (6.8)$$

Esta expresión puede escribirse de la siguiente manera:

$$(E_b/N_0)_j = \frac{W}{\nu_j R_j} \times \frac{P_j}{I_{total} - P_j} \quad (6.9)$$

en la que W es el chiprate, P_j es la potencia recibida del usuario j , ν_j es el factor de actividad del usuario j , R_j es el bitrate del usuario j , e I_{total} es la potencia total recibida incluyendo ruido térmico en la estación base. Despejando P_j se obtiene:

$$P_j = \frac{1}{1 + \frac{W}{(E_b/N_0)_j R_j \nu_j}} \times I_{total} \quad (6.10)$$

Se define la potencia de la manera $P_j = L_j I_{total}$ y se obtiene así el valor de un factor de carga para una conexión.

$$L_j = \frac{1}{1 + \frac{W}{(E_b/N_0)_j R_j \nu_j}} \quad (6.11)$$

La interferencia total recibida, excluyendo el ruido térmico P_N , puede escribirse como la suma de las potencias recibidas por todos los N usuarios que se encuentren en la misma celda.

$$I_{total} - P_N = \sum_{j=1}^N P_j = \sum_{j=1}^N L_j I_{total} \quad (6.12)$$

Para cada nuevo usuario que se incorpora a la red se produce un incremento en el nivel de ruido. Teóricamente este incremento se define como la relación entre la potencia total recibida y la potencia de ruido.

$$N_{rise} = \frac{I_{total}}{P_N} \quad (6.13)$$

Utilizando la expresión 6.12 se obtiene:

$$N_{rise} = \frac{I_{total}}{P_N} = \frac{1}{1 - \sum_{j=1}^N L_j} = \frac{1}{1 - \eta_{UL}} \quad (6.14)$$

donde se ha definido el factor de carga en el uplink, η_{UL} como:

$$\eta_{UL} = \sum_{j=1}^N L_j \quad (6.15)$$

Cuando el factor de carga η_{UL} se aproxima a 1, el incremento de ruido correspondiente se aproxima a infinito y el sistema ha alcanzado su capacidad asintótica N_{pole} .

Por otra parte, en el factor de carga debe tenerse en cuenta la interferencia de otras celdas,

	Definición	Valores
N	Nº usuarios por celda	-
ν_j	Factor de actividad en capa física usuario «j»	0.67 para voz, 1 para datos
E_b/N_0	Relación energía por bit de señal a densidad espectral de ruido	Depende del servicio, bitrate, multicamino, desvanecimientos, canal, antena receptora, velocidad del móvil, etc.
W	Chirate WCDMA	3.84 Mcps
R_j	Bitrate usuario «j»	Depende del servicio
i	Relación de interferencia otras celdas a propia celda visto desde el receptor radio	Macrocelda con antenas omnidireccionales: 55 %. Con 3 sectores: 65 %

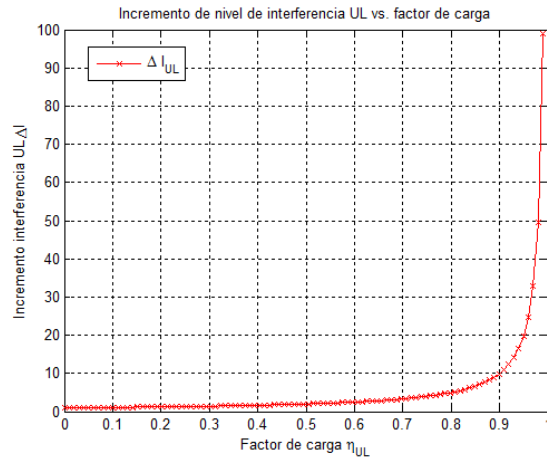
Cuadro 6.5: Definición de parámetros para el cálculo del Factor de Carga Uplink η_{UL} 

Figura 6.21: Incremento de la interferencia UL en función del factor de carga UL

mediante el parámetro i , que representa la relación de interferencia entre otras celdas y la propia celda.

$$i = \frac{Int_{otras,celdas}}{Int_{celda}} \quad (6.16)$$

El factor de carga en el Uplink puede reescribirse como:

$$\eta_{UL} = (1 + i) \sum_{j=1}^N L_j = (1 + i) \sum_{j=1}^N \frac{1}{1 + \frac{W}{(E_b/N_0)_j R_j \nu_j}} \quad (6.17)$$

En la tabla 6.5 se sintetiza la definición de los parámetros que intervienen en las ecuaciones anteriores, así como los valores típicos que adoptan en algunas situaciones.

En la figura 6.21 se representa el incremento de interferencia en función del factor de carga, de acuerdo con la expresión 6.17.

La relación energía por bit de señal a densidad espectral de ruido depende del servicio y de una serie de factores, tal y como refleja la tabla 6.5. Los valores que toma pueden determinarse a partir de simulaciones de nivel de señal, a partir de medidas y a partir de requisitos de desempeño marcados por el 3GPP. La tabla 6.6 muestra algunos valores útiles a la hora del dimensionamiento

Enlace	Tráfico	Valores objetivo
Uplink	CS	5 - 6 dB
	PS	3 - 4 dB
Downlink	CS	6 - 7 dB
	PS	5 - 6 dB
	HSDPA	5 dB (2 Mbps)

Cuadro 6.6: Valores objetivos típicos para la relación E_b/N_0 en el dimensionamiento de redes UMTS

para introducir en las ecuaciones que se han desarrollado en este capítulo hasta el momento. Cabe destacar que los requisitos para tráfico PS son menores, puesto que por su naturaleza es un tráfico que puede ser retransmitido. Por contra, el tráfico CS son datos en tiempo real y no pueden tolerarse retardos, de manera que necesita una relación energía de bit a ruido E_b/N_0 mayor.

6.9. Estudio de la potencia en el Downlink

6.9.1. Factor de carga en el downlink

El factor de carga en el Downlink η_{DL} puede definirse basándose en los mismos principios que para el Uplink, aunque los parámetros de la misma sean ligeramente diferentes, tal y como se representa en la ecuación 6.18.

$$\eta_{DL} = \sum_{j=1}^N \nu_j \frac{(E_b/N_0)_j}{W/R_j} [(1 - \alpha_j) + i_j] \quad (6.18)$$

donde $-10 \log_{10}(1 - \eta_{DL})$ es igual al incremento de ruido sobre el ruido térmico debido a la interferencia por múltiple acceso. Comparada con la ecuación de carga en el uplink, de entre todos los nuevos parámetros que aparecen el más importante es α_j , el cual representa el factor de ortogonalidad en el downlink. WCDMA emplea códigos ortogonales en el downlink para separar a los usuarios y sin ninguna propagación multicamino se mantiene la ortogonalidad cuando la señal de la estación base es recibida por el móvil. No obstante, si hay suficiente dispersión en el retardo en el canal de radio, el móvil verá parte de la señal de la estación base como una interferencia de múltiple acceso. Un valor de ortogonalidad de 1 corresponde a usuarios perfectamente ortogonales. Típicamente, la ortogonalidad se encuentra entre 0.4 y 0.9 en canales multicamino.

En la tabla 6.7 se explica cada uno de los parámetros de la ecuación 6.18 así como sus valores típicos en diferentes situaciones.

En el downlink, la relación de interferencia de otra celda y la propia celda, i depende de la localización del usuario y es, por tanto, diferente para cada usuario j . El factor de carga puede aproximarse por su valor medio a lo largo de la celda, tal y como se muestra en la ecuación 6.19.

$$\overline{\eta_{DL}} = \sum_{j=1}^N \nu_j \frac{(E_b/N_0)_j}{W/R_j} [(1 - \bar{\alpha}) + \bar{i}] \quad (6.19)$$

El factor de carga en el downlink muestra un comportamiento similar a su homólogo en el uplink, en el sentido de que cuando se aproximan a la unidad, el sistema alcanza su capacidad asintótica y el incremento de ruido sobre el ruido térmico tiende a infinito.

	Definición	Valores
N	Nº usuarios por celda	-
ν_j	Factor de actividad en capa física usuario «j»	0.58 para voz, 1 para datos
E_b/N_0	Relación energía por bit de señal a densidad espectral de ruido requerida para alcanzar una BLER predefinida.	Depende del servicio, bitrate, multicamino, desvanecimientos, canal, antena receptora, velocidad del móvil, etc.
W	Chiprate WCDMA	3.84 Mcps
R_j	Bitrate usuario «j»	Depende del servicio
α_j	Ortogonalidad del canal del usuario «j»	Depende de la propagación multicamino. 1: completamente ortogonal, 0: no ortogonal.
i_j	Relación de potencia de otra celda a la propia celda, recibida por el usuario «j».	Cada usuario percibe un valor diferente dependiendo de su posición en la celda y del desvanecimiento log-normal.
$\bar{\alpha}$	Factor de ortogonalidad medio en la celda	Vehículo 50 %, peatón 90 %.
\bar{i}	Relación media entre potencia de la estación procedente de otra celda y de la propia celda recibidas por el usuario.	Macrocelda con antenas omnidireccionales: 55 %. Con tres sectores: 65 %.

Cuadro 6.7: Definición de parámetros para el cálculo del Factor de Carga Uplink η_{UL}

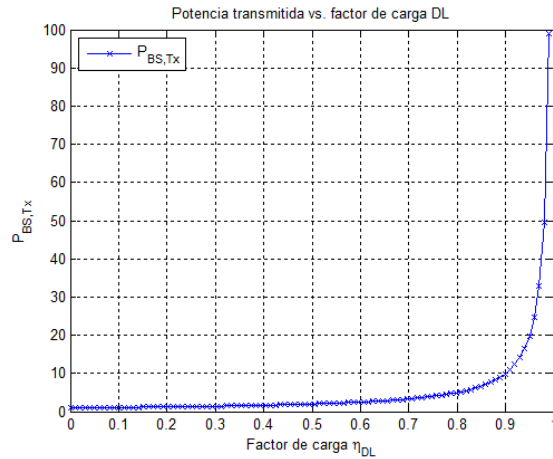


Figura 6.22: Potencia transmitida por el nodo vs. factor de carga DL

6.9.2. Estimación de la potencia total a transmitir

Para el dimensionamiento en el downlink, es importante estimar la potencia total a transmitir por la estación base. Esto debería basarse en la potencia media por usuario y no por la potencia máxima para dar servicio en el límite de la celda a partir de un balance de potencias (*link budget*).

La potencia mínima requerida para la transmisión para cada usuario se determina mediante la atenuación media entre el transmisor de la estación base y el receptor en el equipo de usuario (UE), que se denotará como \bar{L} , y la sensibilidad del receptor en ausencia de interferencia por múltiple acceso (intra o intercelda). Entonces, el efecto del incremento del ruido debido a interferencia se añade a la potencia mínima y el total representa la potencia de transmisión requerida para un usuario situado en una posición intermedia en la celda. Matemáticamente, la potencia total transmitida en el nodo se puede expresar de acuerdo con la ecuación 6.20.

$$BS_{TXP} = \frac{N_{rf} W \bar{L} \sum_{j=1}^N \nu_j \frac{(E_b/N_0)_j}{W/R_j}}{1 - \eta_{DL}} \quad (6.20)$$

En dicha ecuación, N_{rf} es la densidad espectral de ruido del receptor móvil. El valor de N_{rf} puede calcularse con la expresión 6.21 (suponiendo $T=290K$).

$$N_{rf} = kT + NF = -174,0dBm + NF \quad (6.21)$$

donde $k = 1,381 \times 10^{-23} J/K$ es la constante de Boltzmann, T es la temperatura en grados Kelvin y NF es la figura de ruido del receptor del nodo cuyos valores típicos se encuentran en torno a 5-9 dB. En la figura 6.22 se representa una gráfica normalizada que muestra la evolución de la potencia que necesita transmitir el nodo en función del factor de carga.

6.10. Conclusiones

Para que pueda establecerse una conexión UTRAN, deben darse una serie de condiciones:

⇒ Debe haber capacidad en la RNC.

⇒ Debe haber capacidad en el Iub.

- ⇒ Debe haber capacidad de procesamiento banda base en las tarjetas del nodo B (cuya capacidad de procesamiento unitaria se denomina Channel Element).
- ⇒ Debe haber Spreading Code disponibles.
- ⇒ Debe haber potencia en el amplificador del nodo para llegar hasta el User Equipment por encima de su umbral de sensibilidad. Esta potencia será consumida por los diferentes usuarios a los que atiende el nodo, y además dependerá de las pérdidas de propagación hasta alcanzar cada uno de ellos.
- ⇒ En el sentido ascendente, el receptor del Nodo B debe recibir un nivel de interferencia (debida a señal de nodos adyacentes y a la generada por sus propios usuarios) que no le desensibilice, de forma que pueda seguir demodulando la señal uplink del User Equipment.

En este trabajo, sólo se han considerado los recursos susceptibles de ser ampliados por medio de operación de red: capacidad Iub, tarjetas de procesamiento banda base y portadoras radio WCDMA. Se ha descartado la inclusión del resto de factores en el modelado, ya que son recursos sobre los que no hay capacidad de actuación desde el punto de vista de la operación de red.

Sí sería posible reducir la interferencia uplink entre nodos reduciendo el área de solape (soft handover), corrigiendo la inclinación de las antenas y evitando sobrealcances. También se podría reducir la potencia consumida en el amplificador dotando de más nodos a la red, reduciendo así la distancia media entre los usuarios y el nodo (y la atenuación, por consiguiente), y reduciendo el número de usuarios por nodo. Pero estas últimas acciones corresponden a la fase de diseño de la red, no contemplada en este estudio. También queda fuera del alcance de este trabajo la ampliación de capacidad en la RNC.

A lo largo de este capítulo se han caracterizado los principales recursos de que se dispone en las redes UTRAN. De todos ellos, en la presente Tesis son objeto de estudio los recursos:

- ⇒ Capacidad de procesamiento en banda base («Channel Elements»).
- ⇒ Códigos de canalización («Spreading Codes»).
- ⇒ Interfaz Iub.

Así pues, aspectos tales como la interferencia en el enlace ascendente y la potencia en el enlace descendente no serán objeto de estudio. Tras un análisis de la naturaleza de estos factores, se llega a las siguientes conclusiones:

- ⇒ A nivel práctico, es necesario conocer el desglose de usuarios por cada uno de los servicios. Este dato estadístico actualmente no es reportado por los nodos. Teniendo presente que la Tesis utiliza y necesita datos reales para el estudio, no es posible introducir en el modelo desarrollado estos factores, puesto que no existirían datos reales con los que contrastar los cálculos teóricos.
- ⇒ Por otra parte, las configuraciones actuales de las redes UTRAN no da lugar a situaciones en las que estos factores se conviertan en limitantes, ya que una escasa disponibilidad de los mismos a nivel práctico se traduciría también en una escasez de los 3 recursos citados al principio de este apartado y el tráfico se vería bloqueado.

Capítulo 7

Particularización modelo K-R al consumo de recursos

7.1. Introducción

Como ya se ha comentado anteriormente, la red objeto de estudio de esta Tesis soporta múltiples servicios, cada uno de los cuales demanda una cantidad diferente de recursos dentro de la misma. Estos recursos son limitados y es objeto de estudio hacer un dimensionamiento correcto de los mismos para garantizar su correcto funcionamiento y asegurar una determinada calidad de servicio.

El objetivo de este capítulo es describir el proceso de desarrollo del modelo de la red UTRAN propuesto en la presente Tesis. Como se verá posteriormente, se ha modelado un sistema compuesto de varias etapas independientes. El núcleo principal de cálculo de bloqueos es el algoritmo de Kaufman-Roberts en la mayoría de los casos, existiendo otras etapas complementarias que modelan el tráfico y los recursos de forma apropiada.

Cualquier operador de telecomunicaciones necesita conocer el estado de su red. Para ello necesita de mecanismos que le proporcionen el funcionamiento de la misma. Este conocimiento se empleará tanto para detectar o solucionar incidencias en la red como para ser capaces de realizar dimensionamientos acertados de la misma.

El conocimiento del estado de la red se obtiene a través de estadísticos que miden ciertos parámetros como tráfico, velocidades, tiempos de respuesta, carga de procesadores etc. Habitualmente, con estas estadísticas se crean KPIs (*Key Performance Indicators*, o Indicadores Clave de Desempeño), que miden el nivel del desempeño de un proceso, enfocándose en el «cómo» e indicando el rendimiento de los procesos, de forma que se pueda alcanzar el objetivo fijado.

7.2. Planteamiento del problema y aplicación multietapa

Hasta este momento, el estudio de esta Tesis se ha centrado en explicar los fundamentos y el impacto de 5 recursos diferentes, ya comentados en el capítulo 6. No obstante, a la hora de desarrollar el modelo de red UTRAN, dada la limitación tecnológica expuesta en dicho capítulo, no se modelarán todos los recursos. En efecto, si se parte del hecho de que un componente fundamental en el desarrollo de esta Tesis son los datos reales presentados por los nodos, no es posible modelar todos los subsistemas, ya que para modelar ciertos factores, aunque es posible obtener resultados

teóricos, es necesaria la información real, que explícitamente el nodo no reporta detalladamente. Se tienen así datos teóricos en aspectos en los que no se dispone de sus homólogos reales. Así pues, los subsistemas que se modelan (y de los que por tanto son cuantificables sus efectos) son los siguientes:

- ⇒ Capacidad de procesado en banda base (recurso «Channel Elements»).
- ⇒ Recurso códigos de canalización («Spreading Codes»).
- ⇒ Recurso interfaz Iub.

El estudio de la accesibilidad de la red UTRAN deberá tener en cuenta la distinción que existe entre tráfico rígido y de paquetes en el canal Iub, así como la diferenciación entre los enlaces ascendente y descendente. Cada una de estas cuestiones se detallará posteriormente.

Desde un punto de vista funcional, se puede estudiar por separado lo que sucede con cada uno de los recursos. Por ello, el modelo se concibe como una cadena de subsistemas, cada uno de ellos correspondiente a un factor de los 3 planteados. Así, para un determinado tráfico entrante, en función de los recursos que son necesarios para garantizar el servicio y de los que se dispone realmente puede plantearse la aplicación de un algoritmo que permita calcular el bloqueo en cada una de las etapas en serie.

7.2.1. Marco general

Desde un punto de vista general, en un primer nivel de estudio se plantea un proceso como el descrito en la figura 7.1. En este esquema se resume el marco de trabajo del modelo desarrollado, así como su proceso de validación. En la medida de lo posible (tanto en esta figura como en las siguientes) se ha tratado de elaborar una colección de esquemas que permitan al lector comprender de un solo vistazo el funcionamiento del modelo multietapa planteado en esta Tesis. Para favorecer la claridad, se ha decidido presentar al lector diagramas de bloques, que resultan más visuales y con la menor cantidad de texto posible. No obstante, en ocasiones acompañan a los bloques nombres de algunos parámetros para hacer más fácil la comprensión de este apartado por parte del lector.

En el Apéndice A se incluye el código fuente, implementado en Matlab, de cada una de las etapas de que consta el estudio. A lo largo de este capítulo se incluirán referencias a las distintas subrutinas.

El modelo desarrollado será validado con diferentes escenarios reales, Nodos B de una red de un operador real. La información de que se dispone para el estudio está compuesta de:

- ⇒ Estadísticas reportadas por el nodo, tales como intentos de llamadas, peticiones de servicio, etc. En definitiva, información del tráfico cursado por la estación en intervalos de un cuarto de hora para un día completo.
- ⇒ Datos del equipamiento específico con que está dotado el Nodo B o las interfaces Iub. Dicho de otro modo, la capacidad medida en cantidad de recursos disponibles. Normalmente esta información se calcula a partir de las especificaciones de las diferentes tarjetas, sabiendo entonces qué cantidad de recursos posee un determinado nodo.
- ⇒ Valores de los KPIs reales, con los que se podrá evaluar la fiabilidad del modelo.

Como puede verse en la figura 7.1, el modelo, a partir de la información anterior (tráfico y recursos), determina la accesibilidad para el tráfico de paquetes y de circuitos para un periodo

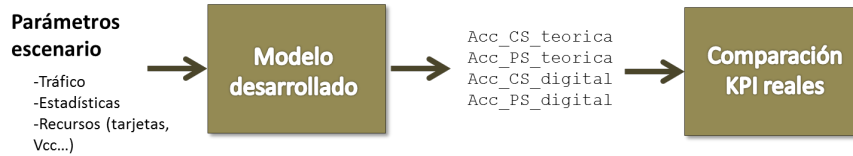


Figura 7.1: Modelo (Primer nivel de desarrollo)

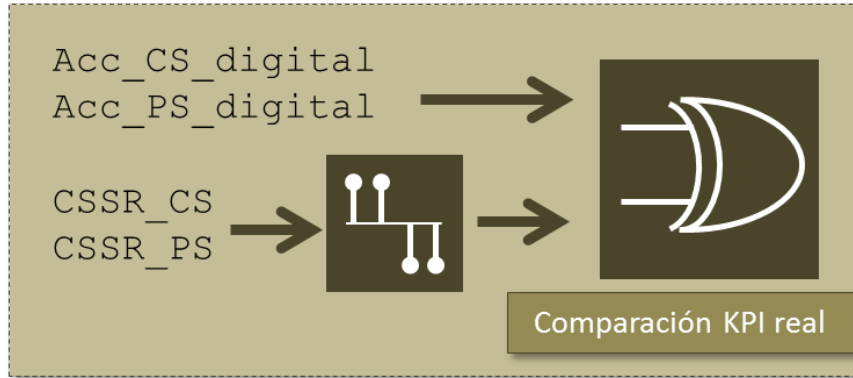


Figura 7.2: Comparación con KPIs reales

de 24 horas. Abstrayendo toda esta información a señales, este resultado puede validarse después desde dos puntos de vista diferentes:

- ⇒ Un enfoque «análogo» permite analizar la fiabilidad del modelo punto a punto comparando los valores absolutos de las accesibilidades teórica y real (recuérdese que la información real viene dada explícitamente de esta manera). En la figura se representan como `Acc_CS_teorica` y `Acc_PS_teorica`. Una comparación desde este punto de vista con el modelo real evalúa por ejemplo el comportamiento global de la accesibilidad cuando se amplía el número de recursos de un nodo o cuando se produce un fallo *hardware*.
- ⇒ Otro punto de vista «digital» lleva a valorar la fiabilidad del modelo basándose en la capacidad que posee de detectar o no la degradación, independientemente de si las curvas de accesibilidades anteriores coinciden. En la figura 7.2 se representa el proceso de validación correspondiente. La información proporcionada por el nodo es siempre «análogica» (`CSSR_CS` y `CSSR_PS`), de manera que es necesario convertirla en una señal binaria («0» significa degradación y «1» no degradación) estableciendo un umbral que defina a partir de qué valor de accesibilidad se considera que existe degradación (en la práctica los operadores asumen que la accesibilidad es buena a partir de 99 %, y así se ha hecho en este modelo). Por otra parte, el modelo desarrollado (fig. 7.1) realiza la correspondiente digitalización de las accesibilidades teóricas (en la figura 7.2 se representan como `Acc_CS_digital` y `Acc_PS_digital`). Ambas parejas de señales digitales son comparadas (`CSSR_CS` con `Acc_CS_digital` y `CSSR_PS` con `Acc_PS_digital`), obteniendo resultados de fiabilidad basadas en la coincidencia o no de diagnósticos de degradación por parte del modelo.

El siguiente nivel de desarrollo del modelo se representa en la figura 7.3. Como se ha comentado anteriormente, será necesario hacer una primera distinción entre enlace descendente (DL) y ascendente (UL), puesto que los factores limitantes no son los mismos (en el *uplink* los códigos de

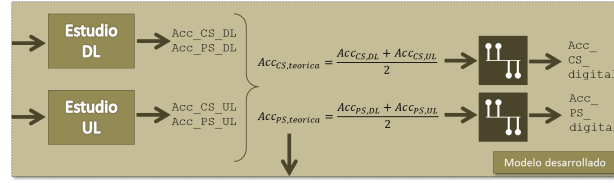


Figura 7.3: Modelo desarrollado (Nivel 2)

canalización no constituyen un factor limitante, ya que cada usuario tiene su propio árbol). Por ello, el modelo se divide en dos subsistemas complementarios, uno para el *downlink* y otro para el *uplink*.

Por otro lado, ha de tenerse en cuenta en el modelado de la red UTRAN que ésta soporta diferentes servicios de forma simultánea, por lo que coexisten en la red diferentes tipos de tráficos. Debido a este motivo, el modelo diferencia entre tráfico rígido o de circuitos (CS) y elástico o de paquetes (PS). Aplicando esta distinción entre CS y PS a cada uno de los subsistemas, se obtienen 4 parámetros de accesibilidad, dos para circuitos (*uplink* y *downlink*) y sus respectivos para paquetes.

Para el posterior proceso de validación hay que tener en cuenta que los datos de accesibilidad proporcionados por el Nodo, si bien distinguen entre tráficos rígido y de paquetes, no existe distinción entre accesibilidad real *downlink* y *uplink*, sino que la información proporcionada combina ambas. Por ello, y para poder establecer comparaciones, es necesario recurrir a algún tipo de métrica que represente lo mismo que los KPIs reales.

Por ejemplo, en el caso del tráfico rígido, definidas las accesibilidades en tantos por ciento, se tiene:

$$Acc_{CS,DL} = \frac{\textit{éxitos CS DL}}{100 \textit{intentos CS DL}} \quad (7.1)$$

$$Acc_{CS,UL} = \frac{\textit{éxitos CS UL}}{100 \textit{intentos CS UL}} \quad (7.2)$$

La métrica buscada representaría los éxitos de tráfico de circuitos con respecto a los intentos totales, es decir:

$$Acc_{CS} = \frac{\textit{éxitos CS}}{\textit{intentos CS}} = \frac{\textit{éxitos CS DL} + \textit{éxitos CS UL}}{200 \textit{intentos CS}} \quad (7.3)$$

De forma análoga, una métrica para medir la accesibilidad de paquetes vendrá dada por la siguiente ecuación:

$$Acc_{PS} = \frac{\textit{éxitos PS}}{\textit{intentos PS}} = \frac{\textit{éxitos PS DL} + \textit{éxitos PS UL}}{200 \textit{intentos PS}} \quad (7.4)$$

Combinando ambas expresiones se obtienen los parámetros $Acc_CS_teorica$ (ecuación 7.3) y $Acc_PS_teorica$ (ecuación 7.4) que son también señales «analógicas» que representan la accesibilidad en función del tiempo para un día completo en intervalos de un cuarto de hora.

Por último, para obtener las expresiones digitales se hace una conversión a binario de las señales anteriores, estableciendo un umbral de comparación o de accesibilidad permitida. Un «0» lógico representa que la accesibilidad en ese momento es inferior al umbral de comparación y un «1» que supera dicho umbral. El umbral de decisión se establecerá en función de cada curva, considerando

que una muestra está degradada cuando su valor difiere en un 1 % del valor medio de la señal en un periodo de 24 horas.

En resumen, el modelo da como resultado cuatro métricas:

- ⇒ Accesibilidad CS (combinando UL y DL) analógica.
- ⇒ Accesibilidad PS (combinando UL y DL) analógica.
- ⇒ Accesibilidad CS digitalizada.
- ⇒ Accesibilidad PS digitalizada.

Estos cuatro parámetros son devueltos por el modelo para su posterior evaluación comparando los resultados con los KPIs reales. El código fuente correspondiente a este marco general puede verse en los apartados A.1.1, A.1.2 y A.2.5.

7.2.2. Modelado del enlace descendente (DL)

El enlace descendente (figura 7.4) se modela como 3 subsistemas en serie que son, ordenadamente:

- ⇒ Subsistema para el recurso «Enlace Iub» (Iub).
- ⇒ Subsistema para el recurso «Channel Elements» (CE).
- ⇒ Subsistema para el recurso «Códigos de Canalización» (SF).

De esta forma, el tráfico entrante en la etapa «Iub» (recurso interfaz Iub), junto con la capacidad del enlace, determina un bloqueo referido a la disponibilidad de este recurso y sus correspondientes accesibilidades CS y PS. El tráfico que ha sido bloqueado se elimina del modelo, de forma que para la etapa siguiente (en ese caso «CE») el tráfico entrante se verá reducido. Con este planteamiento se resuelven los modelos «CE» (recurso Channel Elements) y «SF» (recurso códigos de canalización).

Por último, las accesibilidades teóricas referidas a cada subsistema han de ser combinadas, obteniéndose así unos valores de accesibilidad que engloban las distintas posibilidades de bloqueo que pueden darse en el sistema (en la figura 7.4 Acc_CS_DL y Acc_PS_DL). En los apartados sucesivos se detallan más ampliamente estos subsistemas.

7.2.2.1. Subsistema Iub (DL)

Este subsistema modela el enlace Iub desde el punto de vista de tráfico *downlink*. El recurso limitante será pues la capacidad de este enlace, donde los diferentes tipos de tráfico competirán por un determinado ancho de banda que se verá limitado.

En este subsistema debe hacerse distinción entre el tráfico rígido o de circuitos (CS) y el tráfico elástico o de paquetes (PS), de forma que el estudio de accesibilidad se desdoblará en 2 cálculos paralelos. En la figura 7.5 se representa este subsistema mediante bloques para una configuración *dual stack*.

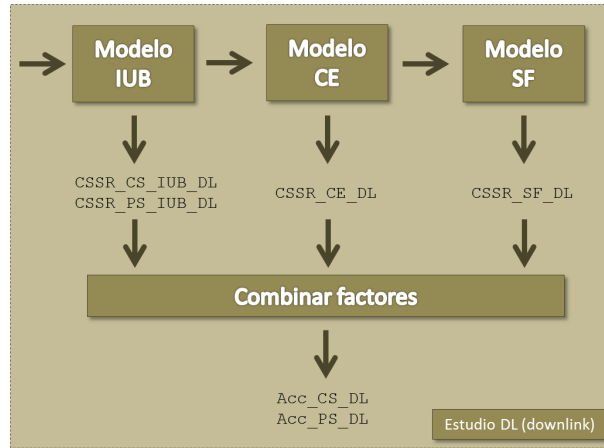


Figura 7.4: Modelo del enlace descendente (DL)

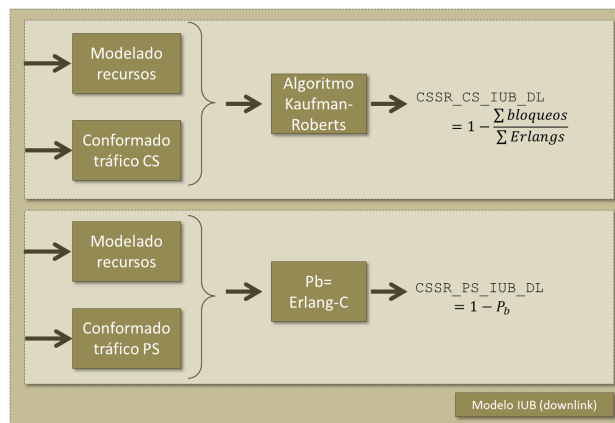


Figura 7.5: Modelo IUB en el downlink

Accesibilidad tráfico de circuitos

La parte superior de la figura muestra el tratamiento que se da al tráfico de circuitos. El núcleo del estudio es el algoritmo de Kaufman-Roberts, que permite, como ya se ha visto en capítulos anteriores, el cálculo del bloqueo de cada uno de los servicios en una red multiservicio como es la red UTRAN. Se necesita conocer por tanto:

- ⇒ El tráfico (en Erlang) entrante, desglosado para cada tipo de servicio. Existe un subsistema específico para el conformado de tráfico que se explicará posteriormente.
- ⇒ La demanda unitaria (ancho de banda en Kbps) que requiere cada tipo de tráfico.
- ⇒ La capacidad total disponible por la que «compiten» todos los tráficos. Se modela a partir del equipamiento del nodo. Posteriormente se detallará cómo se realiza el modelado de los recursos.

Con esta información, aplicando el algoritmo de Kaufman-Roberts, se obtiene la probabilidad de bloqueo para cada tipo de tráfico, y conocido el tráfico entrante, es posible determinar el nº de llamadas bloqueadas por el sistema.

La accesibilidad de circuitos ($CSSR_{CS,Iub,DL}$) se calculará mediante la expresión 7.5:

$$CSSR_{CS,Iub,DL} = 1 - \frac{\sum \text{bloqueos}}{\sum \text{Erlangs}} \quad (7.5)$$

Accesibilidad tráfico de paquetes

La accesibilidad Iub de paquetes en el enlace descendente (parte inferior de la figura 7.5), donde se transmite a través de IP, se calcula aplicando la fórmula de Erlang C, ya que en este caso se tiene en cuenta que si una petición llega cuando todo está ocupado, ésta se introduce en una cola a la espera de poder ser atendida. Aunque el modelo teórico asume que la cola tiene capacidad infinita, en la realidad la red la implementa con capacidad finita, de manera que cuando dicha cola está completa la siguiente llamada se bloquea provocando rechazos. Una aplicación sucesiva de esta fórmula en cada instante de tiempo permite obtener otro vector $CSSR_{PS,Iub,DL}$ que representa la accesibilidad de tráfico de paquetes en la interfaz Iub en el *downlink*.

La ecuación 7.6 representa el cálculo de dicha accesibilidad. P_b es pues un vector que contiene, en cada elemento $P_b[n]$ el bloqueo calculado según la fórmula de Erlang-C para un instante de tiempo.

$$CSSR_{PS,Iub,DL} = 1 - P_b \quad (7.6)$$

Nodos configurados como *Full-IP*

Si el nodo se configura como *Full IP*, el cálculo se realiza de forma similar. La diferencia radica en que la capacidad total se comparte para circuitos y paquetes, al transmitirse todo por IP. La capacidad disponible para circuitos se obtiene de crear un Vcc virtual, de 1.272 Mbps. La capacidad restante, una vez ocupada la de circuitos, es la que hay disponible para paquetes (igual para el enlace ascendente y el descendente). Una vez conocida la capacidad de la que se dispone, se realizan los mismos pasos y cálculos que en el caso de nodos con transmisión *DualStack*: es decir, se calcula la accesibilidad de circuitos utilizando Kaufman-Roberts y la accesibilidad de paquetes con la fórmula de Erlang C.

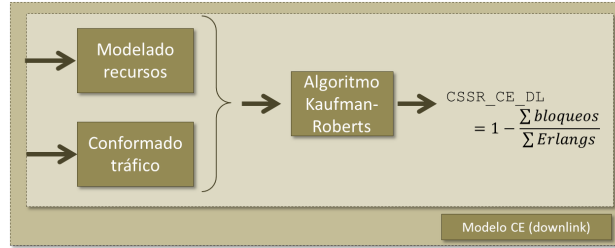


Figura 7.6: Modelo CE en el downlink

El código fuente que implementa el estudio de la accesibilidad del recurso Iub se encuentra en los apartados A.2.1.1 para nodos Dual Stack y A.2.1.2-A.2.1.2 para nodos Full-IP.

7.2.2.2. Subsistema CE (DL)

Este subsistema modela el recurso Channel Elements, que representa la capacidad de procesamiento en banda base que tiene el Nodo B. En este subsistema no se hace distinción explícita entre circuitos y paquetes. Debe tenerse en cuenta que dado que se trata de subsistemas ordenados (ver figura 7.4), el *tráfico*¹ de entrada en este subsistema no es el *tráfico* inicial. Si existe bloqueo, como previsiblemente sucederá en algún instante, en el subsistema Iub, el *tráfico* entrante en este segundo subsistema será sólo una fracción del inicial, de manera que siendo $Att_{iub,DL}$ los intentos de llamada (*tráfico* entrante) al subsistema anterior y $Acc_{iub,DL}$ la accesibilidad calculada, los intentos de llamada a considerar en la segunda etapa vendrán dados por:

$$Att_{CE,DL} = Att_{iub,DL} \cdot Acc_{iub,DL} \quad (7.7)$$

donde $Att_{CE,DL}$ representa en cada instante un vector de intentos con los diferentes tipos de tráfico de cada servicio que se ofrecen en ese momento. En los apartados A.4.2.1 y A.4.2.2 se implementa este procedimiento.

En la figura 7.6 se muestra detalladamente el subsistema que estudia el bloqueo para el recurso Channel Elements. Al igual que en el caso anterior, el núcleo de cálculo es de nuevo el algoritmo multiservicio de Kaufman-Roberts con los mismos requisitos que los comentados anteriormente. Por ello, previa ejecución del algoritmo K-R se efectúa un pre-proceso modelando el tráfico y los recursos como en el caso anterior.

De esta manera, la accesibilidad vendrá dada por la ecuación 7.8, definida como el complementario de la probabilidad de bloqueo (fallos respecto a los intentos).

$$CSSR_{CE,DL} = 1 - \frac{\sum \text{bloques}}{\sum \text{Erlangs}} \quad (7.8)$$

El código fuente que implementa el estudio de la accesibilidad para el recurso Channel Elements puede encontrarse en el apartado A.2.2.

¹El término «Tráfico» se utiliza aquí de manera general siguiendo la nomenclatura de modelos de colas que se presentan en la bibliografía. En realidad, cuando se plantea un esquema multietapa, este término no se corresponde con Erlangs sino con intentos de conexión.

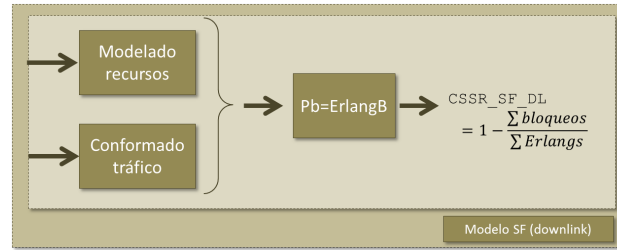


Figura 7.7: Modelo SF en el downlink

7.2.2.3. Subsistema SF (DL)

En la figura 7.7 se representa el modelo para el recurso de códigos de canalización (Spreading Codes) en el *downlink*. El bloqueo se calcula mediante la fórmula de Erlang-B, ya que no todos los servicios consumen recursos del mismo tipo. Al igual que en etapas anteriores, el tráfico y los recursos serán modelados de acuerdo a respectivos subsistemas que se explicarán posteriormente.

La accesibilidad parcial obtenida en este caso vendrá dada por la ecuación 7.9.

$$CSSR_{SF,DL} = 1 - \frac{\sum \text{bloqueos}}{\sum \text{Erlangs}} \quad (7.9)$$

El código fuente que implementa el estudio del recurso códigos de canalización puede encontrarse en el apartado A.2.3, y la fórmula de Erlang-B se implementa en el apartado A.3.8.

7.2.3. Combinación de factores en DL (IUB/CE/SF)

Al comienzo de esta sección se ha presentado un modelo para el enlace descendente describiéndolo como una sucesión de subsistemas, cada uno de los cuales se dedicaba al estudio del bloqueo existente para un determinado recurso. Se comentó además que, abstrayendo cada uno de los subsistemas adecuadamente, es posible obtener probabilidades de bloqueo (y por tanto accesibilidades) para cada uno de los modelos.

Una vez realizados estos cálculos, es necesario determinar una métrica que englobe el comportamiento del tráfico a nivel general, incluyendo cada uno de los factores (canal Iub, CE y SF). Durante el desarrollo de la Tesis se realizaron dos enfoques en este sentido:

⇒ Modelo de bloques.

⇒ Modelo probabilístico.

7.2.3.1. Modelo de bloques

Una manera de determinar la accesibilidad total a partir de las accesibilidades parciales es utilizando un modelo de bloques, en el que cada etapa representa un subsistema y se tienen sendos flujos de entrada y de salida. El enfoque es el mismo, tanto si se trata de tráfico rígido como elástico. En la figura 7.8 se representa el cálculo para el tráfico de circuitos (voz, videollamada y señalización).

El desarrollo teórico que se presenta a continuación es válido tanto para tráfico CS (voz, videollamada y señalización) como PS (paquetes PS384 y HSPA), ya que como se verá a continuación, la accesibilidad no dependerá del tráfico de entrada. Para ambos supuestos, sea $\rho_{i,IUB,DL}$ el tráfico de entrada al sistema.

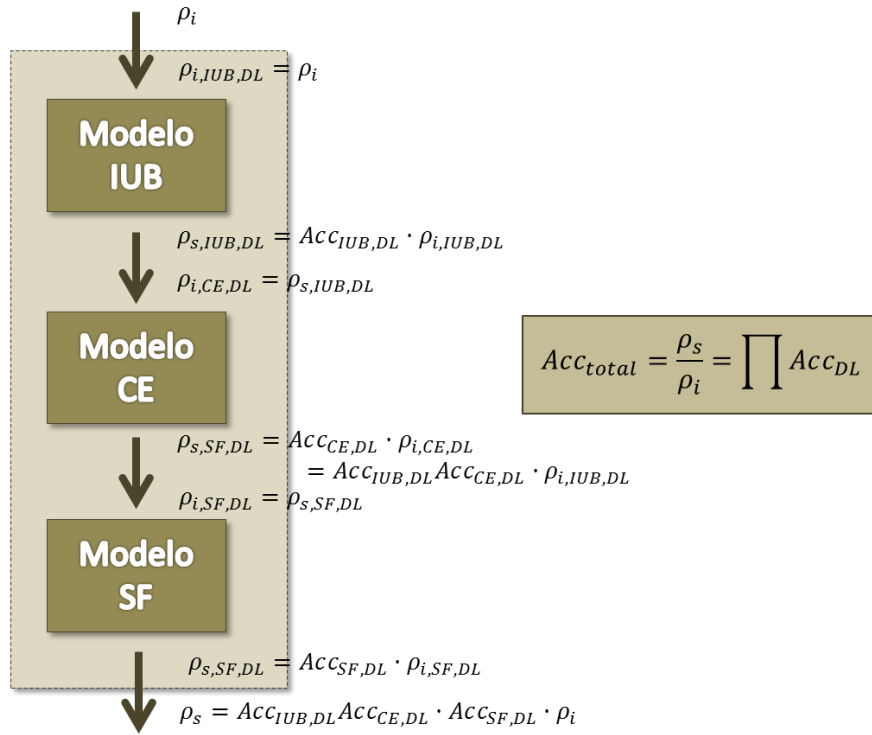


Figura 7.8: Modelo de bloques para calcular la accesibilidad total CS

Subsistema	Tráfico entrante (ρ_i)	Accesibilidad (Acc)	Tráfico saliente (ρ_s)
IUB	$\rho_i = \rho_{i,IUB,DL}$	$Acc_{IUB,DL}$	$\rho_i \cdot Acc_{IUB,DL}$
CE	$\rho_i \cdot Acc_{IUB,DL}$	$Acc_{CE,DL}$	$\rho_i \cdot Acc_{IUB,DL} \cdot Acc_{CE,DL}$
SC	$\rho_i \cdot Acc_{IUB,DL} \cdot Acc_{CE,DL}$	$Acc_{SC,DL}$	$\rho_i \cdot Acc_{IUB,DL} \cdot Acc_{CE,DL} \cdot Acc_{SC,DL}$

Cuadro 7.1: Tráficos entrante y saliente en modelo downlink

Teniendo en cuenta que el orden de los subsistemas en el enlace descendente es IUB-CE-SC, se definen:

A partir de la tabla anterior, teniendo en cuenta que lo que se ha definido como tráfico saliente del subsistema SC representan los éxitos de conexión, se define la accesibilidad global del sistema como:

$$Acc_{Total} = \frac{\text{éxitos}}{\text{intentos}} = \frac{\rho_{i,IUB,DL} \cdot Acc_{IUB,DL} \cdot Acc_{CE,DL} \cdot Acc_{SC,DL}}{\rho_{i,IUB,DL}} \quad (7.10)$$

$$Acc_{Total} = Acc_{IUB,DL} \cdot Acc_{CE,DL} \cdot Acc_{SC,DL}$$

de manera que la accesibilidad total es el producto de las accesibilidades parciales de cada uno de los subsistemas, no influyendo en ningún caso la definición que se haga para el tráfico de entrada. No obstante, y dado que el factor $Acc_{IUB,DL}$ es diferente para tráfico CS que para PS, se tienen dos valores de accesibilidad total.

7.2.3.2. Modelo probabilístico

Otro posible enfoque que lleva a deducir la misma expresión es considerando el bloqueo de la red UTRAN dentro de un espacio probabilístico, tal y como se representa en la figura 7.9, donde

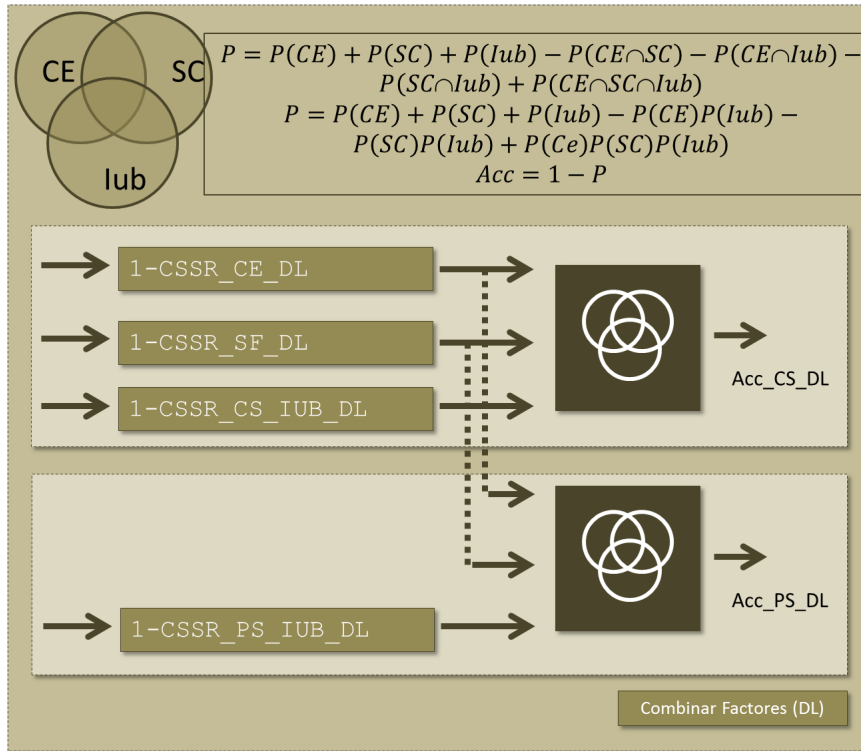


Figura 7.9: Combinación de factores en el DL.

se incluye además el proceso de combinación, independiente del modelo elegido.

Considérese un espacio muestral Ω derivado de un espacio probabilístico (Ω, β, P) . Considerando β el conjunto de los sucesos:

- ⇒ Bloqueo en el subsistema CE («Channel Elements»).
- ⇒ Bloqueo en el subsistema SC («Spreading Codes»).
- ⇒ Bloqueo en la interfaz Iub.

se define la probabilidad P de los sucesos como:

- ⇒ $P(CE)$, probabilidad de bloqueo en el subsistema CE.
- ⇒ $P(SC)$, probabilidad de bloqueo en el subsistema SC.
- ⇒ $P(IUB)$, probabilidad de bloqueo en el interfaz Iub.

La probabilidad de que una conexión se vea bloqueada vendrá dada por la expresión :

$$P_B = P(CE) + P(SC) + P(IUB) - P(CE \cap SC) - P(CE \cap IUB) - P(SC \cap IUB) + P(CE \cap SC \cap IUB) \quad (7.11)$$

Teniendo en cuenta que el bloqueo en cada uno de los subsistemas es independiente, la expresión anterior se puede reescribir como:

$$P_b = P(CE) + P(SC) + P(IUB) - P(CE)P(SC) - P(CE)P(IUB) - P(SC)P(IUB) + P(CE)P(SC)P(IUB) \quad (7.12)$$

Y por tanto, la accesibilidad combinando los tres factores vendrá dada según la ecuación 7.13:

$$Acc_{CS/PS,DL} = 1 - P_b \quad (7.13)$$

Desarrollando las expresiones de bloqueo parciales como las complementarias de las correspondientes accesibilidades se concluye que las expresiones 7.10 y 7.13 son iguales.

Teniendo en cuenta que el subsistema Iub considera por separado el tráfico de circuitos y de paquetes, se obtienen dos valores de accesibilidad correspondientes a ambos tipos de tráfico:

- ⇒ Accesibilidad de circuitos combinada DL (Acc_{CS_DL}). Se aplicarán las expresiones anteriores, teniendo en cuenta que C es la probabilidad de bloqueo del recurso Iub para el tráfico de circuitos.
- ⇒ Accesibilidad de paquetes combinada DL (Acc_{PS_DL}). En este caso C es el bloqueo del recurso Iub para el tráfico de paquetes.

En el apartado A.2.4 se encuentra la implementación de las expresiones anteriores.

7.3. Modelado del enlace ascendente

La figura 7.10 representa el modelo que se ha desarrollado para el enlace ascendente. A diferencia del downlink explicado anteriormente, este modelo presenta las siguientes diferencias:

- ⇒ No se implementa el estudio del recurso códigos de canalización, ya que no constituye un factor limitante. Como se ha comentado en el capítulo anterior, cada usuario dispone de su propio árbol de códigos, de manera que no es un recurso compartido.
- ⇒ El orden de los subsistemas es el inverso, primero se analiza el recurso channel elements y después el tráfico que no haya sido bloqueado se analiza en el modelo del Iub UL.

Así pues, se obtienen en este caso 3 parámetros diferentes:

- ⇒ El subsistema que estudia los CE determina una accesibilidad referida a este recurso ($CSSR_{CE_UL}$).
- ⇒ El modelo Iub en el UL distingue entre tráfico de circuitos y de paquetes, con lo que se calculan sendas accesibilidades CS y PS ($CSSR_{CS_IUB_UL}$ y $CSSR_{PS_IUB_UL}$).

Por último, y para poder concluir accesibilidad de paquetes y de circuitos, es necesario combinar las accesibilidades parciales de cada uno de estos subsistemas, tal y como se explicará a continuación. Tras esta combinación se obtienen:

- ⇒ Accesibilidad de circuitos CS (Acc_{CS_UL}).
- ⇒ Accesibilidad de paquetes PS (Acc_{PS_UL}).

7.3.1. Subsistema CE (UL)

En la figura 7.11 se presenta el subsistema encargado de estudiar la accesibilidad correspondiente al recurso Channel Elements en el enlace ascendente.

El núcleo principal de cálculo es el algoritmo de Kaufman-Roberts, con todas las consideraciones hechas previamente, para lo cual será necesario preparar el tráfico y modelar los recursos para que sea consistente.

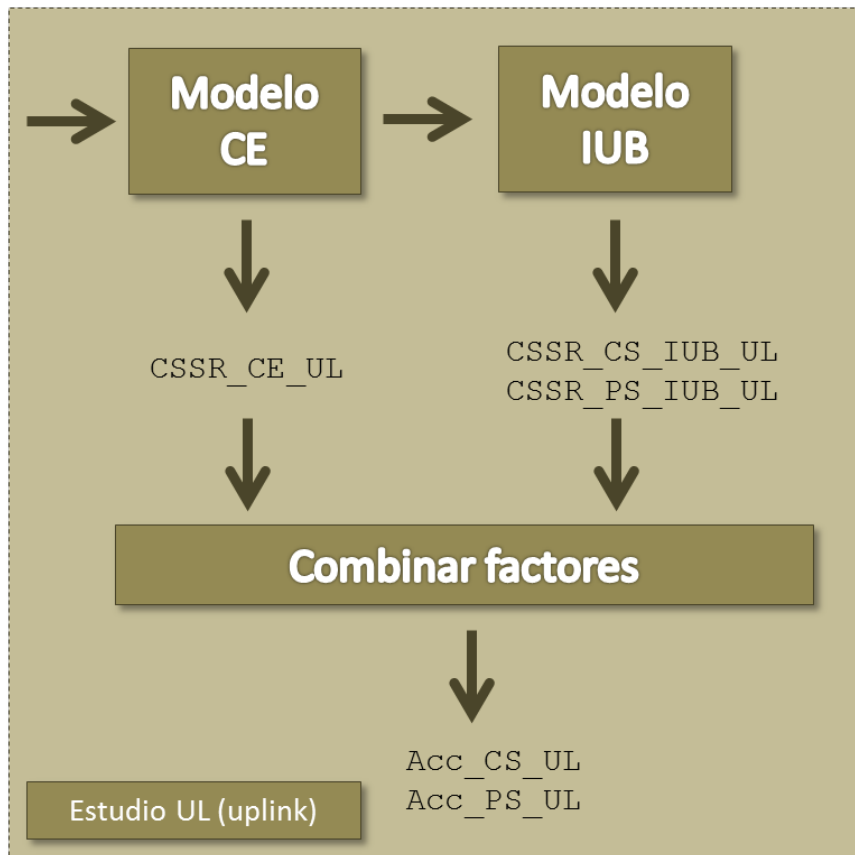


Figura 7.10: Modelo para el enlace ascendente.

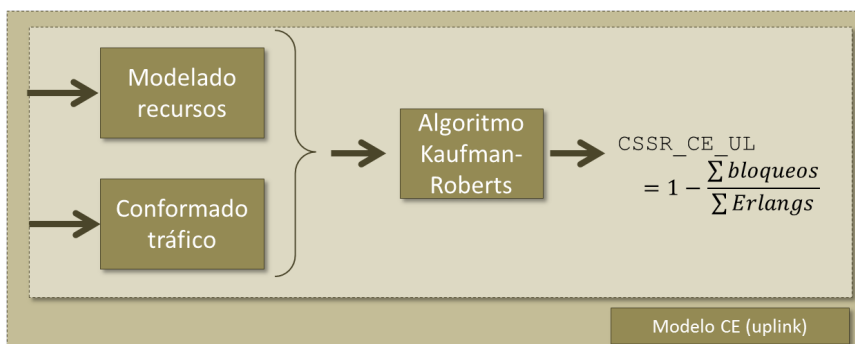


Figura 7.11: Modelo CE (UL)

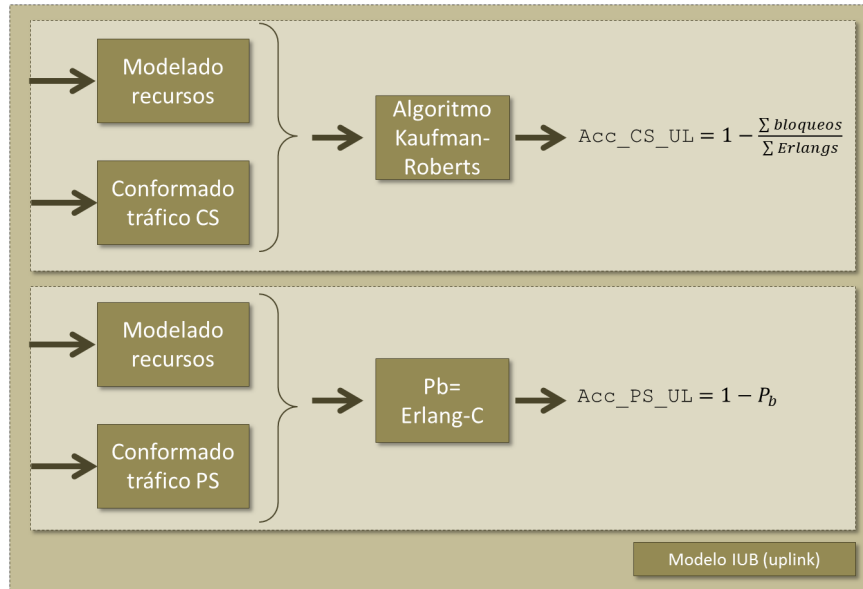


Figura 7.12: Modelo Iub (UL)

Este modelo calcula la accesibilidad de acuerdo con la expresión 7.14:

$$CSSR_{CE,DL} = 1 - \frac{\sum \text{bloqueos}}{\sum \text{Erlangs}} \quad (7.14)$$

La función implementada en el apartado A.2.2 sirve también para el estudio del enlace ascendente.

7.3.2. Subsistema Iub (UL)

Este subsistema modela el enlace Iub en sentido ascendente (ver figura 7.12). Como en el caso del enlace DL, se estudia por separado el tráfico de circuitos y de paquetes:

- ⇒ La accesibilidad relativa al tráfico CS (Acc_CS_UL) se determina según el algoritmo de Kaufman-Roberts, para el cálculo de bloqueo en redes multiservicio. Todas las consideraciones realizadas anteriormente en relación a este algoritmo son de aplicación también en este punto.
- ⇒ La relativa al tráfico PS (Acc_PS_UL) se calcula mediante la fórmula de Erlang-C, teniendo en cuenta que la cola de espera es de longitud finita y por tanto genera bloqueo. En este caso no se transmite mediante IP si no a través de ATM, por lo que se comparte la capacidad total que aportan los E1 entre los servicios de circuitos y de paquetes, como se explicará posteriormente.

En el apartado A.2.1.1 se encuentra la implementación en Matlab del recurso Channel Elements, utilizado indistintamente para el enlace ascendente y descendente. Para el tráfico de paquetes, se utiliza la función Erlang-C, implementada en el apartado A.3.9.

Subsistema	Tráfico entrante	Accesibilidad	Tráfico saliente
CE	$\rho_{i,CE,UL}$	$Acc_{CE,UL}$	$\rho_{i,CE,UL} \cdot Acc_{CE,UL}$
IUB	$\rho_{i,CE,UL} \cdot Acc_{CE,UL}$	$Acc_{IUB,UL}$	$\rho_{i,CE,UL} \cdot Acc_{CE,UL} \cdot Acc_{IUB,DL}$

Cuadro 7.2: Tráficos en modelo UL

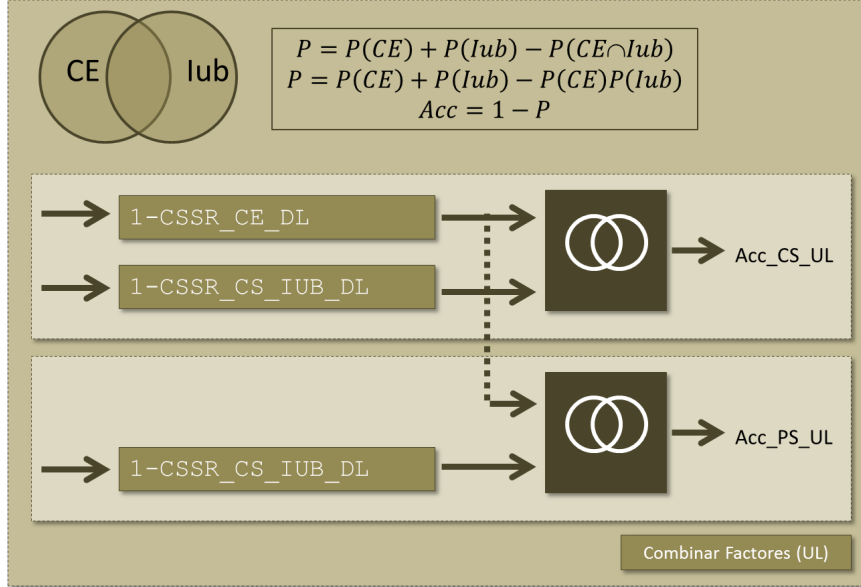


Figura 7.13: Combinación de factores UL

7.3.3. Combinación de factores en UL (CE/Iub)

La obtención de una accesibilidad global, combinación de las accesibilidades parciales determinadas anteriormente para el enlace ascendente sigue la misma metodología que en el enlace descendente.

7.3.3.1. Modelo de bloques

El planteamiento es el mismo que se ha comentado anteriormente para el downlink. Teniendo en cuenta que el orden de los subsistemas en el enlace ascendente es CE-IUB, se definen:

A partir de la tabla anterior, teniendo en cuenta que lo que se ha definido como tráfico saliente del subsistema IUB representan los éxitos de conexión, se define la accesibilidad global del sistema como:

$$Acc_{Total} = \frac{\text{éxitos}}{\text{intentos}} = \frac{Att_{CE,UL} \cdot Acc_{IUB,UL} \cdot Acc_{IUB,UL}}{Att_{CE,UL}} = Acc_{IUB,UL} \cdot Acc_{IUB,UL} \quad (7.15)$$

7.3.3.2. Modelo probabilístico

El esquema ahora (ver figura 7.13) se simplifica al tener en cuenta que sólo existen dos factores limitantes (recurso CE y recurso Iub).

Considérese un espacio muestral Ω derivado de un espacio probabilístico (Ω, β, P) . Considerando β el conjunto de los sucesos:

\Rightarrow Bloqueo en el subsistema CE («Channel Elements»).

⇒ Bloqueo en la interfaz Iub.

se define la probabilidad P de los sucesos como:

⇒ $P(CE)$, probabilidad de bloqueo en el subsistema CE.

⇒ $P(IUB)$, probabilidad de bloqueo en la interfaz Iub.

La probabilidad de que una conexión se vea bloqueada vendrá dada por la expresión :

$$P_B = P(CE) + P(IUB) - P(CE \cap IUB) \quad (7.16)$$

Teniendo en cuenta que el bloqueo en cada uno de los subsistemas es independiente, la expresión anterior se puede reescribir como:

$$P_b = P(CE) + P(IUB) - P(CE)P(IUB) \quad (7.17)$$

Y por tanto, la accesibilidad combinando los tres factores vendrá dada según la ecuación 7.18:

$$Acc_{CS/PS,DL} = 1 - P_b \quad (7.18)$$

Dado que en el modelo UL se incluye el subsistema Iub, y puesto que en el mismo se hace distinción entre tráfico de circuitos y de paquetes, se consideran sendas accesibilidades:

⇒ Accesibilidad de circuitos combinada UL (Acc_{CS_UL}). Se obtiene con las expresiones anteriores siendo B el bloqueo del recurso Iub para el tráfico de circuitos.

⇒ Accesibilidad de paquetes combinada UL (Acc_{PS_UL}), donde B es el bloqueo del recurso Iub para el tráfico de paquetes.

La combinación de estos dos factores se implementa en la misma función utilizada para el enlace descendente, que puede encontrarse en el apartado A.2.4.

7.4. Conformado de tráfico

Como se ha comentado, el nodo B almacena estadísticas de tráfico cada 15 minutos, representando cada muestra el acumulado de los 15 minutos previos. Por ejemplo, considerando intentos de llamada, la primera muestra del día representa el n^o total de intentos entre las 0:00 y las 0:15, la segunda muestra los intentos desde las 0:15 hasta las 0:30, y así sucesivamente. Dado que el tráfico objeto de estudio presentará grandes fluctuaciones en intervalos de tiempo mucho más pequeños, se hace necesaria una reducción de la resolución temporal (por ejemplo un valor ideal y aceptable sería 1 minuto). Sin embargo, según el funcionamiento actual de los equipos instalados en las redes UMTS, el fabricante no proporciona de forma real estadísticas con menor resolución. A pesar de ello, es posible estimar cómo se distribuirá el tráfico a lo largo de esos 15 minutos de forma teórica recurriendo a distribuciones de probabilidad, lo que implicará interpolar las muestras reportadas.

Se trata pues de repartir cómo se distribuye ese agregado de tráfico que representa cada muestra original en 15 nuevas muestras: en definitiva, estimar la distribución del tráfico minuto a minuto. Para ello, es necesario determinar la distribución de probabilidad y sus parámetros correspondientes, tales como su media o su varianza. Para la elección de la distribución que mejor se adapte a los escenarios que serán objeto de estudio, es imprescindible tener en cuenta las siguientes consideraciones:

7.4.1. Forma y parámetros de la distribución

La presente Tesis se fundamenta en el modelado de peticiones de servicio que se producen a ráfagas llegando a saturar el sistema, por ejemplo un incremento sustancial del número de llamadas entrantes durante un breve intervalo de tiempo (por ejemplo, el descanso de un evento deportivo). Por ello, se debe pensar en una distribución cuya densidad tenga forma de campana o similar, y que además su varianza sea pequeña para concentrar en la medida de lo posible el tráfico en intervalos cortos.

7.4.1.1. Enfoque aproximativo desde el punto de vista infinitesimal

Recurriendo a cualquier texto de estadística matemática una variable aleatoria x puede ser definida mediante su función de densidad de probabilidad $f(x)$, con $x \in \mathbb{R}$, cuya característica fundamental aquí es que el área bajo la curva en toda la recta real es igual a la unidad (expresión 7.19).

$$\int_{-\infty}^{\infty} f(x) = 1 \quad (7.19)$$

De esta manera, un valor K puede teóricamente descomponerse/repartirse a lo largo de la recta real sin más que multiplicarlo por la curva de densidad de obabilidad $f(x)$, tal y como muestra la expresión 7.20. Este paso puede interpretarse como un escalado de la curva de densidad por una constante K , originándose una nueva curva de densidad escalada $g(x)$ con $x \in \mathbb{R}$.

$$g(x) = K \cdot f(x) \quad (7.20)$$

Así, puede generalizarse que la muestra original K queda repartida a lo largo de la «ventana» o del dominio $x \in [-\infty, +\infty] = \mathbb{R}$ en infinitas muestras. La reconstrucción de la misma vendrá dada por la ecuación 7.21.

$$K = K \cdot \int_{-\infty}^{\infty} f(x) \quad (7.21)$$

En el caso de que una cualquiera de las curvas $f(x)$ o $g(x)$ fuera enventanada en un dominio finito el valor del área bajo la curva ya no sería la unidad. En la figura 7.14 se representa la curva de una distribución de Poisson enventanada en el dominio $x \in [0, 15]$. La ecuación 7.21 se transforma entonces en una integral definida, tal y como indica la expresión 7.22. El valor de la muestra original queda pues mermado en una cantidad más o menos significativa en función del área de la curva que «caiga» fuera de la ventana. La expresión 7.23 define el error de enventanado, que queda representado en la figura 7.15. El comportamiento de esta gráfica se explica por sí mismo: a medida que aumenta el valor de λ el valor medio de la curva se va desplazando hacia valores crecientes de x y su varianza también aumenta haciendo la curva más ancha. En consecuencia, una parte cada vez más significativa del flanco derecho de la curva «cae» fuera de la ventana, con lo que el error de reconstrucción posterior aumenta.

$$K_0 = K \cdot \int_0^{15} f(x) < K \quad (7.22)$$

$$\epsilon = K - K_0 \quad (7.23)$$

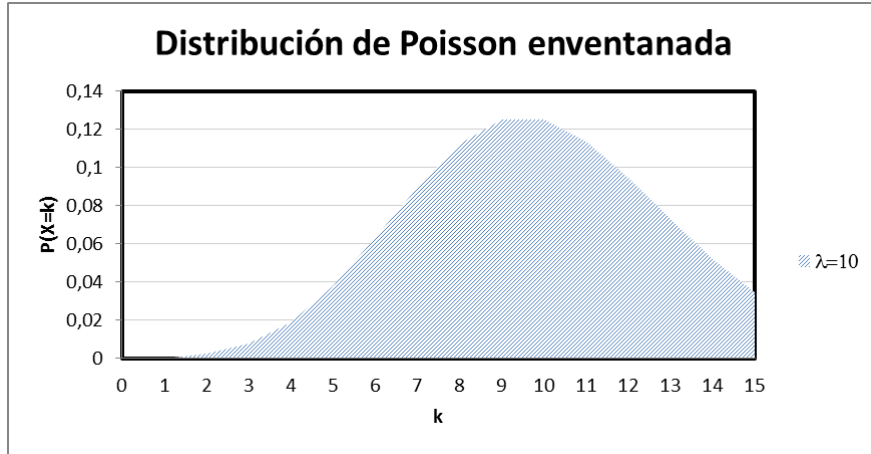


Figura 7.14: Problema del enventanado de la función de densidad de probabilidad para una distribución no uniforme

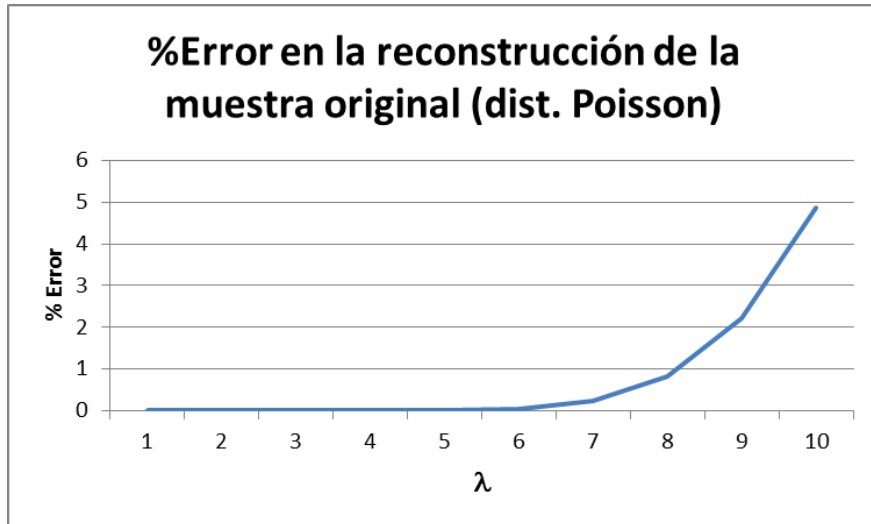


Figura 7.15: Error en la reconstrucción de la muestra original para una distribución de Poisson enventanada

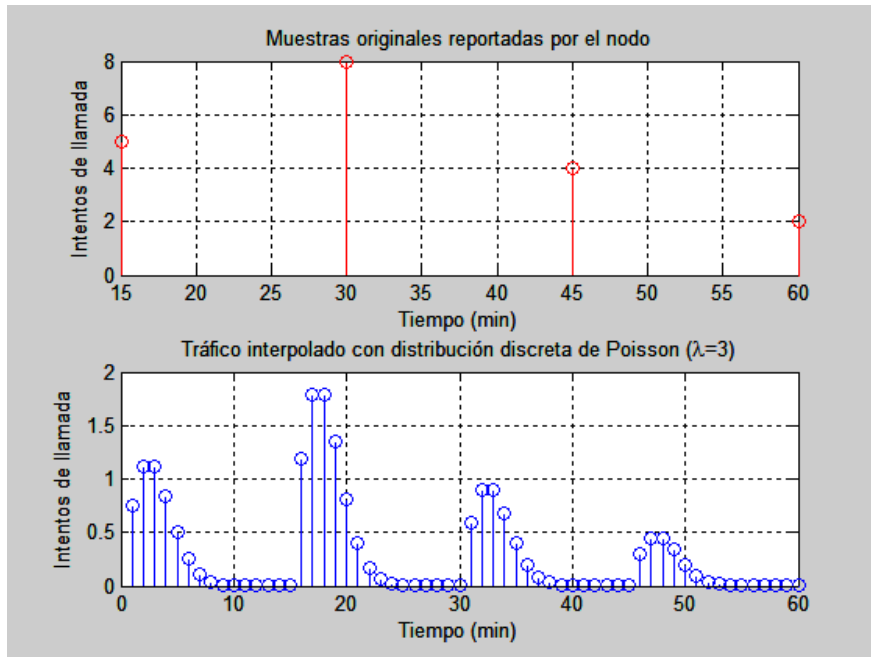


Figura 7.16: Ejemplo de interpolación de tráfico

7.4.1.2. Particularización práctica desde el punto de vista discreto

La descomposición previa de una muestra K en infinitas muestras tiene únicamente valor teórico. Sin embargo, el problema real del reparto difiere del primer planteamiento en dos puntos:

- ⇒ Debe definirse en un dominio de longitud finita, acotando los límites superior e inferior: $x \in [0, 15]$.
- ⇒ La muestra original K debe repartirse en un n° finito: 15 muestras, por lo que la variable aleatoria x se transforma en una variable aleatoria discreta n con $n \in [0, 15]$, $\forall n \in \mathbb{N}$.

Con estas observaciones, el problema anterior transforma las curvas de densidad de probabilidad $f(x)$ y $g(x)$ en sendos vectores de interpolación $f[n]$ y $g[n]$, con $n \in \mathbb{N}$, que no son si no las curvas de densidad de probabilidad elegidas ahora muestreadas. La operación de reparto se puede definir por tanto según la expresión 7.24. La figura 7.16 representa una hipotética situación, en la que se representan los intentos de llamada para los primeros 60 minutos. La gráfica superior muestra los datos reportados por el nodo, mientras que en la inferior se representa el tráfico interpolado $g[n]$ con un vector de reparto $x[n]$ de acuerdo a una distribución de Poisson de parámetro $\lambda = 3$ (se ha elegido un valor que facilite la representación gráfica para hacerlo más legible al lector).

Al igual que antes, el resultado de ponderar la muestra de valor K con una ventana de longitud finita da como resultado un valor $K_0 < K$. Posteriormente, el proceso de interpolación deberá ser revocado reconstruyendo el vector original de muestras. En consecuencia, puede definirse nuevamente un parámetro ϵ como error de enventanado o de reconstrucción, como diferencia entre el valor original de la muestra interpolada y el agregado real suma de las muestras ponderadas (ecuación 7.25).

$$K_0 = \sum_{n=0}^{15} g[n] < K \quad (7.24)$$

$$\epsilon = K - K_0 \quad (7.25)$$

Este proceso de reparto provoca inevitablemente un error de reconstrucción, con lo que el objetivo en este punto debe ser la elección de vectores de ponderación adecuados.

- ⇒ La media de la distribución deberá estar comprendida en el intervalo $n \in [0, 15]$, $\forall n \in \mathbb{N}$. Una media muy próxima a los extremos provoca la caída de un flanco fuera de la ventana con una pérdida muy significativa en la reconstrucción de la muestra original.
- ⇒ La varianza deberá ser pequeña (corroborando lo ya comentado en el apartado anterior), ya que curvas muy anchas también provocan la caída de parte de la curva fuera de la ventana.

7.4.1.3. Elección del vector de reparto (función de distribución discreta)

La elección de cómo repartir un valor agregado K , en un número finito de muestras plantea, como se ha descrito anteriormente, multitud de posibilidades. Así, en principio es posible definir una variable aleatoria discreta con cualquier curva de densidad de probabilidad y así la ponderación será pues función de la elección de dicha curva. Tal y como se ha comentado en el apartado anterior, es deseable que cumpla una serie de requisitos que se pueden concretar en dos:

- ⇒ El tráfico debe quedar repartido de forma coherente.
- ⇒ La distribución deberá tener unos parámetros adecuados (media, varianza...) que permitan reconstruir las muestras iniciales con el mínimo error de enventanado posible, ya que cuanto mayor sea el valor de ϵ se estará considerando menos tráfico del que realmente existe, luego las consecuencias pueden diferir de la situación real.

Para la elección de la distribución que seguirá el vector de reparto se estudiarán distintas opciones:

- ⇒ Reparto equitativo o distribución uniforme.
- ⇒ Distribución normal o gaussiana.
- ⇒ Distribución de Poisson.

7.4.1.4. Reparto equitativo o distribución uniforme

Una primera consideración puede ser el reparto del valor agregado con igual probabilidad a lo largo del tiempo, modelando por tanto el tráfico con una función de densidad de probabilidad uniforme. Este modelado presenta como principal ventaja que es absolutamente preciso, ya que no deja lugar a la variación de ningún parámetro (media, varianza, etc). Sin embargo, este reparto no modela el tráfico según las características objeto de estudio, ya que no se presentan picos de tráfico.

7.4.1.5. Distribución normal

Considerando que el tráfico debe repartirse de manera que dentro de cada intervalo el tráfico presente un incremento significativo, existen varias distribuciones que pueden estudiarse. Este es precisamente el punto fuerte de modelar el tráfico con esta distribución. Sin embargo, como inconveniente se hará necesario manejar 2 parámetros distintos e independientes, su media y su varianza, lo cual le añade cierta complejidad al estudio.

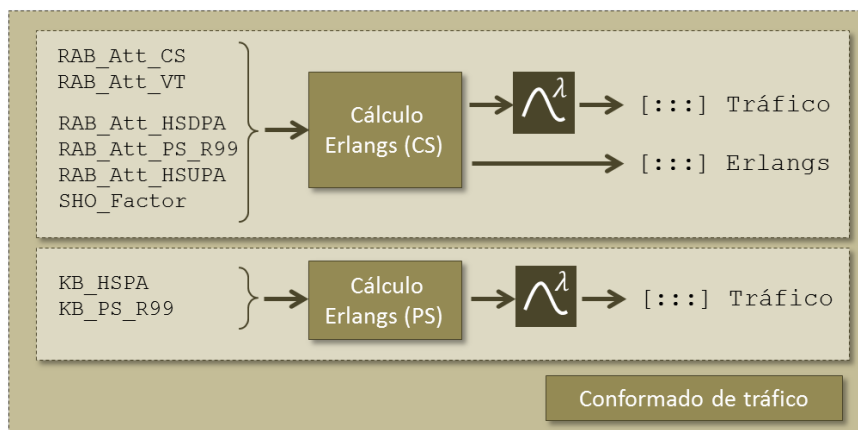


Figura 7.17: Subsistema de conformado de tráfico

7.4.1.6. Distribución de Poisson

Existe una amplia bibliografía con la que se puede argumentar a favor de elegir como distribución de tráfico una distribución de Poisson. La llegada de clientes a un sistema, o lo que es lo mismo, de peticiones de servicio a la red móvil, es un proceso que puede ser modelado como proceso de Poisson con una determinada tasa de llegadas λ . Como ventajas con respecto a la distribución normal pueden citarse:

- ⇒ Tiene un respaldo teórico más fuerte. No obstante, es necesario tener presente que se trata de modelos teóricos y que su coincidencia con el comportamiento real tendrá indiscutiblemente un cierto error.
- ⇒ La función de densidad de probabilidad de la distribución de Poisson viene determinada con un único parámetro λ , que es al mismo tiempo su media y su varianza. Se reduce por tanto un grado de libertad que simplifica sustancialmente el modelado del tráfico.

7.4.2. Implementación del conformado de tráfico en el modelo UTRAN

Tal y como se ha comentado en la sección anterior, la aplicación del modelo de Kaufman-Roberts requiere, entre otros parámetros, el tráfico de entrada. Para modular en la medida de lo posible cada una de estas funciones, en el modelo desarrollado se ha optado por abstraer la aplicación del algoritmo K-R en un subsistema y preparar los datos que éste necesita en otros subsistemas complementarios. En este apartado se describe el subsistema de conformado de tráfico (ver figura 7.17), cuya implementación se encuentra en los apartados A.3.3 y A.3.4. Además, la función que realiza la posterior reconstrucción del tráfico para recuperar una resolución temporal de 15 minutos se encuentra en el apartado A.4.3.

7.4.2.1. Cálculo del tráfico en Erlangs²

En la tabla 7.3 se muestra los tipos de tráfico de entrada que aplican para cada subsistema.

²Debe tenerse en cuenta que la información proporcionada por el nodo a nivel de estadísticas ha de tener en cuenta no solamente las peticiones de servicio cursadas por el mismo, sino también el tráfico procedente del *soft handover* de las celdas vecinas.

	Iub		CE		SF
	DL	UL	DL	UL	DL
Voz	X	X	X	X	X
Videollamada	X	X	X	X	X
Señalización	X				
PS384	·		X	X	X
HSDPA	·		X		X
HSUPA	·			X	

Cuadro 7.3: Tráficos CS de entrada utilizados en cada subsistema

	Servicio	Erlangs IUB	Erlangs CE/SC
CS	Voz	$\rho_{voz} = min_{voz}/15$	
	VT	$\rho_{VT} = min_{vt}/15$	
	Señalización	$\rho_{sign} = \rho_{voz} + \rho_{VT}$	$\rho_{sign} = 0$
PS	PS384	$\rho_{ps384} = \lambda/\mu$	$\rho_{ps384} = RAB_Att_PS_R99$
	HSDPA	$\rho_{hsdpa} = \lambda/\mu$	$\rho_{hsdpa} = RAB_Att_HSDPA$
	HSUPA	$\rho_{hsupa} = \lambda/\mu$	$\rho_{hsupa} = RAB_Att_HSUPA$

Cuadro 7.4: Cálculo de tráfico en Erlangs

Se supondrá, para este modelo de tráfico, que durante el periodo de observación, las conexiones existentes se mantienen en curso. A partir de los datos reportados por el nodo es necesario operar con ellos para obtener el valor del tráfico de entrada para el modelo:

⇒ Tráfico CS (voz y videollamada (+señalización en el IUB)). Considerando que 1 Erlang se define como la ocupación de un canal o conexión, dado que el tráfico se reporta como el agregado de los últimos 15 minutos, el tráfico en Erlangs para estos servicios se define como los minutos totales dividido por los minutos de observación, es decir $\rho_{voz/VT} = min/15$, siendo *min* los minutos de voz o de videollamada. En el enlace IUB entre el nodo y la RNC debe considerarse también el tráfico debido a la señalización. Cada Erlang de voz o de videollamada lleva asociado otro Erlang de señalización.

⇒ Tráfico PS (paquetes PS384 y HSPA) se define de manera diferente en función del subsistema que se considere:

- En el subsistema IUB se define como $\rho = \lambda/\mu$, siendo λ el régimen binario (bps) y μ un valor que puede ser elegido arbitrariamente (por ejemplo 1 Mbps). Para determinar el valor de λ , dado que el nodo reporta KB y recordando que ese valor es el agregado de los últimos 15 minutos, es inmediato obtener «bps» sin más que aplicar el factor:

$$\lambda(bps) = KB \cdot \frac{8}{900}$$

- En los subsistemas CE y SC el concepto es diferente, ya que considerando que 1 Erlang representa un recurso ocupado (en este caso árboles o channel elements), puede considerarse que cada intento de conexión es un Erlang.

De este modo, el tráfico en Erlangs estará relacionado con los intentos de RAB segun la tabla 7.4:

7.4.2.2. Implementación del tráfico

Ya se ha comentado que la resolución ofrecida por el nodo B a través de sus estadísticas resulta insuficiente, ya que registra datos acumulativos cada 15 minutos. Para poder disminuir artificialmente esta resolución, es necesario realizar una interpolación del tráfico, cuyas particularidades ya han sido descritas con anterioridad. Como se puede ver en la figura 7.17, el subsistema de conformado de tráfico implementado definitivamente realiza una interpolación según una distribución de Poisson de parámetro $\lambda = 1$.

Así pues, y recordando que el propósito del subsistema de conformado de tráfico es preparar éste para ser introducido como parámetro de entrada a posteriores subsistemas de cálculo de bloqueo (véase algoritmo de Kaufman-Roberts), se obtienen fruto de este módulo dos matrices (expresiones 7.26 y 7.27):

$$\overline{\overline{[E]}} = \begin{pmatrix} \overrightarrow{E_1} \\ \overrightarrow{E_2} \\ \dots \\ \overrightarrow{E_n} \end{pmatrix} \quad (7.26)$$

$$\overline{\overline{[\rho]}} = \begin{pmatrix} \overrightarrow{\rho_1} \\ \overrightarrow{\rho_2} \\ \dots \\ \overrightarrow{\rho_n} \end{pmatrix} \quad (7.27)$$

Ambas matrices están definidas de la siguiente manera:

- ⇒ La matriz $[E]$ representa el tráfico calculado en Erlangs (datos cada 15 minutos) y la matriz $[\rho]$ el tráfico interpolado según la distribución de Poisson (tráfico minuto a minuto). (En la figura 7.17, matrices **Erlangs** y **Tráfico**, respectivamente).
- ⇒ En ambos casos, cada fila de la matriz representa un tipo de tráfico (por ejemplo, $\overrightarrow{E_1}$ puede representar tráfico de voz $\overrightarrow{E_2}$ tráfico de videollamada, etc).
- ⇒ Cada una de esas filas es un vector que representa el tráfico elemento a elemento a lo largo del tiempo (por ejemplo, E_{11} puede representar el tráfico de voz en el primer cuarto de hora, E_{12} en el segundo, etc).

7.5. Modelado de recursos

El contexto en el que se plantea el modelo de la Tesis está basado en el cálculo del bloqueo, producido cuando un sistema de capacidad finita se ve infradimensionado: una nueva petición de servicio demanda una cantidad de recursos que el sistema no es capaz de satisfacer, por lo que la llamada o la petición se bloquean.

El modelado de los recursos es necesario a la hora de implementar cualquier cálculo de bloqueo. Por ejemplo, recuerde el lector que el algoritmo de Kaufman-Roberts necesita como parámetros de entrada:

- ⇒ El tráfico entrante, ya descrito en el apartado anterior.
- ⇒ Los recursos de que se dispone, generalmente limitados, y por los que competirá ese tráfico.

	Iub (kbps)		CE		SF
	DL	UL	DL	UL	DL
Voz	12	12	1	1	256
Videotelefonía	64	64	4	2	32
Señalización	6	-	-	-	-
PS384	-	-	16	8	32
HSDPA	-	-	1.3	-	16
HSUPA	-	-	-	2	2/4

Cuadro 7.5: Modelado de recursos

⇒ La capacidad total disponible.

En este apartado se describe el segundo punto, correspondiente al modelado de recursos en la red UTRAN en base a los 3 factores considerados en los cálculos a nivel práctico: Iub, CE y SF. La tabla 7.5 recoge la demanda unitaria para cada tipo de tráfico.

Una vez que están claros cuáles son los tipos de tráfico que intervienen en cada subsistema (ver apartado anterior), así como la demanda unitaria de cada uno de ellos (tabla 7.5), se construye un vector:

$$\vec{b} = \left(b_1 \quad b_2 \quad \dots \quad b_n \right) \quad (7.28)$$

formado por todas las demandas unitarias de los servicios. Por ejemplo, para el enlace Iub, b_1 puede representar la demanda unitaria en kbps del servicio voz, b_2 del servicio de videotelefonía, y así sucesivamente.

La implementación del modelado de recursos se encuentra en el apartado A.3.1 para el tráfico CS y A.3.2 para el tráfico PS.

7.6. Cálculo de la capacidad del sistema

7.6.1. Capacidad del Iub

En este caso, y dado que el estudio de la accesibilidad se hace por separado para tráficos CS y PS, les corresponderán a ambos respectivamente sendas capacidades.

7.6.1.1. Nodos Dual-Stack

Enlace descendente

Para el tráfico CS, la capacidad de la interfaz Iub es la capacidad de cada VCC medida en Kbps por el número total de VCC que haya en el escenario. La capacidad del VCC se calcula teniendo en cuenta que hay 3000 celdas ATM por segundo, y cada celda tiene 53 bytes, dando como resultado una capacidad de 1.272 Mbps en cada VCC.

Para la accesibilidad de paquetes, ya se ha comentado anteriormente que la transmisión se realiza a través de IP, con lo que la capacidad del Iub será la capacidad que tenga dicha conexión IP.

Enlace ascendente

Para el tráfico de circuitos, la capacidad Iub se determina igual que en el downlink, teniendo en cuenta el número total de VCC.

Sin embargo, para el tráfico de paquetes no se transmite mediante IP si no a través de ATM, por lo que se comparte la capacidad total que aportan los E1 entre los servicios de circuitos y de paquetes. Una vez se ha reservado la capacidad necesaria para circuitos, la capacidad disponible para paquetes es la sobrante. Así, se calcula la capacidad de los servicios de paquetes de forma que, si se parte de un número determinado de E1, cada uno con una capacidad de 2.048 Mbps, hay que restar a esa capacidad la que se reserva para los Vcc en circuitos para obtener la utilizada en paquetes.

Por ejemplo, en un enlace con 5 E1 en el que se reservan 2 Vcc para circuitos, la capacidad disponible para el enlace ascendente en paquetes sería:

$$5 \times 2.048 \text{ Mbps} - 2 \times 1.272 \text{ Mbps} = 7.696 \text{ Mbps}$$

7.6.1.2. Nodos Full-IP

El cálculo se realiza de forma similar. La diferencia radica en que la capacidad total se comparte para circuitos y paquetes, al transmitirse todo por IP. La capacidad disponible para circuitos se obtiene de crear un Vcc virtual, de 1.272 Mbps. La capacidad restante, una vez ocupada la de circuitos, es la que hay disponible para paquetes. Esto es igual para el enlace ascendente y el descendente.

7.6.2. Capacidad para el recurso Códigos de Canalización (SF)

La capacidad, se calcula el número de códigos disponibles que hay de SF determinado. Para obtenerla, se parte de dos árboles de códigos de spreading completos, es decir, cada uno con:

⇒ 1 código de SF 1

⇒ 2 códigos de SF 2

⇒ 4 códigos de SF 4

⇒ 8 códigos de SF 8

⇒ 16 códigos de SF 16

⇒ 32 códigos de SF 32

⇒ 64 códigos de SF 64

⇒ 128 códigos de SF 128

⇒ 256 códigos de SF 256

Como existen dos arboles, al final se tiene el doble de códigos de cada SF, es decir, inicialmente se parte de un total de:

⇒ 2 código de SF 1

⇒ 4 códigos de SF 2

- ⇒ 8 códigos de SF 4
- ⇒ 16 códigos de SF 8
- ⇒ 32 códigos de SF 16
- ⇒ 64 códigos de SF 32
- ⇒ 128 códigos de SF 64
- ⇒ 256 códigos de SF 128
- ⇒ 512 códigos de SF 256

Inicialmente se hace una reserva de códigos de SF16 para el servicio de HSDPA. Se pueden reservar 5, 10 o 15 códigos. Permitir que HSDPA opere con 15 códigos no resulta muy realista en la mayoría de los casos, porque ocuparía casi todo el árbol y no se podrían atender eficientemente otros servicios. Asignar 5 o 10 códigos ofrece un razonable equilibrio entre el rendimiento para el usuario individual y la capacidad de la célula. En este caso se opta por reservar inicialmente 10 códigos de cada árbol para HSDPA, es decir 20 códigos de SF16 en total. El primer paso, una vez reservados, es calcular el número de códigos que quedan libres para el resto de servicios. A su vez se puede calcular la probabilidad de bloqueo para HSDPA a través de la fórmula de Erlang-B, teniendo en cuenta que la capacidad es de 20 y de que el número de peticiones es conocido.

Una vez conocido la cantidad de códigos que quedan libres, se procederá a ocupar los necesarios para dar cabida a otro, además de calcular su probabilidad de bloqueo utilizando la fórmula de Erlang-B, y así sucesivamente con todos los servicios. Una vez calculada la probabilidad de bloqueo de cada servicio, se puede obtener la accesibilidad como el inverso de esa probabilidad, y así calcular la accesibilidad total por códigos de spreading como el producto de las accesibilidades de cada servicio.

7.6.2.1. Detalles de implementación

En la figura 7.19 se muestra de forma esquemática el proceso para el cálculo del bloqueo en este contexto, donde el factor limitante son los códigos de canalización.

7.6.3. Capacidad para el recurso Channel Elements (CE)

7.6.3.1. Enlace descendente

La forma de calcular la accesibilidad por Channel Elements para el enlace descendente y ascendente es similar, solo cambia la cantidad de CE de los que se disponen, así como el consumo de CE por parte de los servicios. El primer cálculo que se realiza es para obtener la capacidad total de Channel Elements que se tienen. En los escenarios analizados en esta Tesis, para el enlace descendente se utilizan la tarjetas HSTX45 y HSTX60, que tienen las siguientes características:

HSTX45

- ⇒ Se parte de 384 CE.
- ⇒ Si se activa HSUPA se consumen 128 CE.
- ⇒ Si se activa el segundo ResourceId se consumen otros 128 CE.

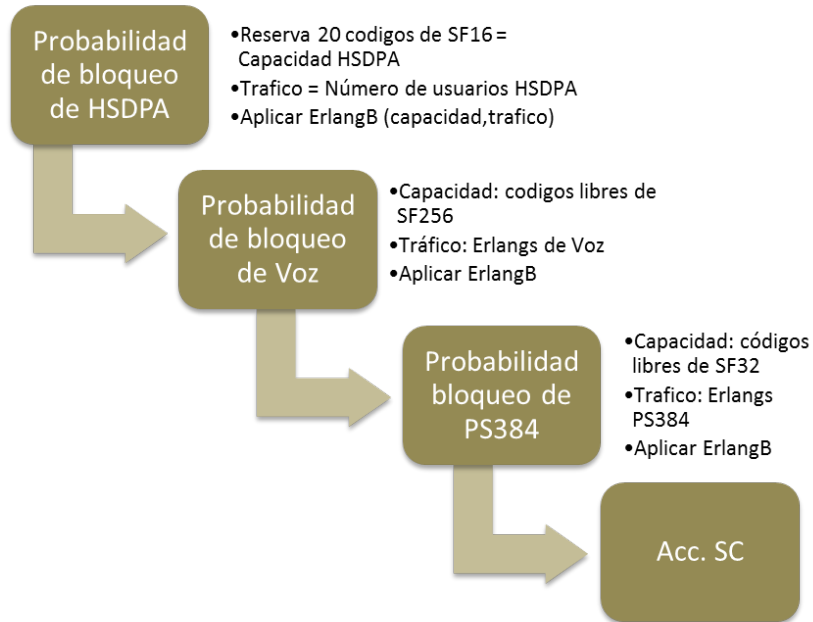


Figura 7.18: Esquema de cálculo de la accesibilidad por códigos de Spreading

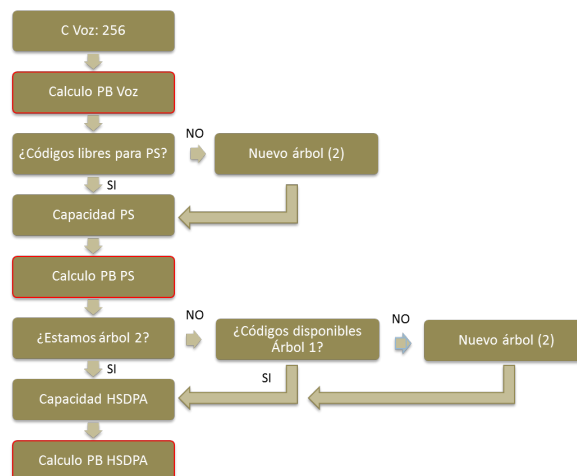


Figura 7.19: Esquema para el cálculo del bloqueo debido a códigos de canalización

⇒ Si se activa el tercer ResourceId se consumen 128 CE.

⇒ Por cada 16 usuarios de HSDPA se consumen 21 CE.

HSTX60

⇒ Se parte de 384 CE.

⇒ Por defecto está activado HSUPA.

⇒ Si se activa el segundo ResourceId no se consumen CE.

⇒ Si se activa el tercer ResourceId se consumen 128 CE.

⇒ Por cada 16 usuarios de HSDPA se consumen 21 CE.

En este caso se tiene activada la opción de HSUPA y el segundo ResourceId. Si el escenario a estudiar tiene, al menos, una tarjeta del tipo HSTX60, se activarán en ella el HSUPA y el segundo ResourceId porque no implica ningún gasto adicional de CE, con lo que sería la resolución óptima. Si por el contrario, sólo hay tarjetas del modelo HSTX45, se deberá activar el HSUPA y el segundo ResourceId en una de ellas, consumiendo 128 CE por cada activación.

Por tanto, la capacidad total se calcula de la siguiente forma:

Si hay una o más tarjetas HSTX60 (independientemente del número de tarjetas HSTX45):

$$C_{total} = N_{HSTX45} \times C_{HSTX45} + N_{HSTX60} \times C_{HSTX60} \quad (7.29)$$

Si no hay ninguna tarjeta HSTX60 (sólo hay HSTX45):

$$C_{total} = [(N_{HSTX45} - 1) \times C_{HSTX45}] + [C_{HSTX45} - 128 - 128] \quad (7.30)$$

siendo, $C_{HSTX45} = C_{HSTX60} = 384$ y N_{HSTX45} y N_{HSTX60} el número de tarjetas HSTX45 y HSTX60, respectivamente.

7.6.3.2. Enlace ascendente

Para el enlace ascendente, la capacidad de CE que se tiene es la capacidad inicial que tienen las tarjetas. Para esta Tesis, se utilizan los modelos RAX13 y RAX14, los cuales tienen 128 CE.

$$C_{total} = N_{RAX13} \times C_{RAX13} + N_{RAX14} \times C_{RAX14} \quad (7.31)$$

siendo, $C_{RAX13} = C_{RAX14} = 128$ y N_{RAX13} y N_{RAX14} el número de tarjetas RAX13 y RAX14, respectivamente.

7.7. Cálculo del bloqueo

Uno de los bloques fundamentales del modelo desarrollado en esta Tesis es el correspondiente al cálculo de la probabilidad de bloqueo (y en última instancia, de la accesibilidad como complemento). Como ya se ha descrito en los capítulos precedentes, la naturaleza de la red UTRAN exige (salvo casos particulares descritos en este capítulo) la aplicación de un algoritmo capaz de calcular el

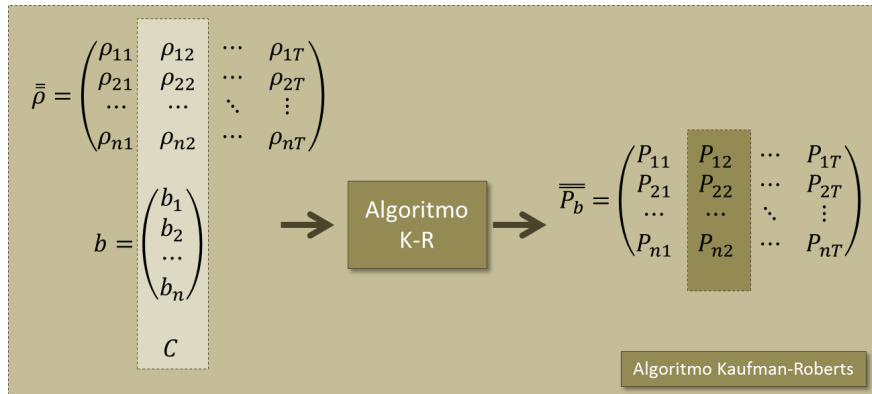


Figura 7.20: Subsistema cálculo de bloqueo para el algoritmo K-R

bloqueo de una red multiservicio. A lo largo de la Tesis se ha descrito el algoritmo de Kaufman-Roberts y se han discutido detalles de su implementación. El lector si así lo desea puede recurrir al apéndice correspondiente para consultar la implementación en código fuente.

El objetivo de este apartado es poner de manifiesto la correspondencia entre la aplicación explícita del algoritmo de Kaufman-Roberts y los subsistemas complementarios que preparan los datos para su introducción, así como la información que se espera obtener tras la ejecución del mismo.

Recordando una vez más cuáles son los parámetros del algoritmo de Kaufman-Roberts, se tiene:

- ⇒ Tráfico de entrada. Deberán indicarse los Erlangs de tráfico para cada uno de los servicios. Ya se ha comentado anteriormente que el tráfico queda definido en forma matricial según las expresiones 7.26 y 7.27.
- ⇒ Demanda unitaria. Representa el n° de unidades de recurso que necesita cada tipo de tráfico. Los recursos quedan dispuestos según un vector (expresión 7.28).
- ⇒ Capacidad del sistema. Es el n° de unidades de recurso que existen en total, y por las que competirán los diferentes tipos de tráfico. Es un valor constante y se define de acuerdo con el apartado anterior.

A partir de la información anterior, el bloqueo queda calculado en una nueva matriz $[P_b]$ que tiene, en cada columna, las probabilidades de bloqueo de los diferentes servicios en un instante de tiempo particular.

En las figuras precedentes, a la hora de describir los diferentes modelos (Iub CS, CE, SF...) se mostró el contexto en el cual se ubicaba el cálculo del bloqueo. En la figura 7.20 se muestra de manera esquemática la metodología con la que se aplica el algoritmo K-R (subsistema «Algoritmo de Kaufman-Roberts» en figuras precedentes).

En dicha figura se ilustra la manera en la que se calcula la matriz de probabilidades de bloqueo $[P_b]$, para T instantes de tiempo repartidos a lo largo del día.

- ⇒ Se toma una columna de la matriz de tráfico, que contiene el tráfico en Erlangs para n servicios. En la figura se ha sombreado como ejemplo los tráfico correspondientes al minuto 2.
- ⇒ Se utilizan como constantes el vector \vec{b} , que contiene la demanda unitaria de unidades de recurso y C , capacidad total del sistema.

- ⇒ Con esta información, se aplica el algoritmo de Kaufman-Roberts, obteniéndose un vector que contiene las probabilidades de bloqueo para cada uno de los n servicios.
- ⇒ Este resultado se anexa como una nueva columna en la matriz $[P_b]$. En el ejemplo de la figura 7.20, al tratarse del tráfico en el instante $t=2$, el bloqueo resultante se sitúa en la segunda columna.
- ⇒ El estudio continúa con el tráfico en el instante siguiente (en el ejemplo sería $t=3$, o tercera columna de $[\rho]$, y así sucesivamente hasta completar todos los instantes de tiempo.

Este procedimiento puede aplicarse de forma análoga si el algoritmo de bloqueo es otro distinto al de K-R. En los apartados A.3.5, A.3.6 y A.3.7.

7.8. Planteamiento de implementación del algoritmo de K-R para otros posibles factores

7.8.1. Implementación del algoritmo K-R para el análisis de la interferencia

Como ya es sabido, el número de usuarios que puede atender un nodo B de forma simultánea es limitado. Entre otros factores, porque a medida que crece la demanda de tráfico disminuye la relación señal a interferencia. Así, por debajo de un determinado umbral, el Nodo B no es capaz de demodular correctamente la información recibida y en consecuencia la comunicación se ve degradada.

Este margen de trabajo se va estrechando a medida que crece la demanda de clientes y los servicios requieren una mayor tasa binaria.

Se ha definido en el capítulo anterior el concepto de «factor de carga». Se comprueba si existe interferencia analizando el factor de carga UL, de forma que si este factor es mayor que 1 se puede asegurar que existe interferencia. Lo deseable es obtener un factor de carga menor que 1. Para realizar el cálculo del factor de carga es necesario conocer el valor de los diferentes parámetros que afectan al mismo, algunos de los cuales depende del tipo de servicio, y que ya han sido explicados con detalle en páginas anteriores.

Con todo ello se puede calcular el factor de carga UL y comprobar si hay interferencias o no. Existe dos opciones:

- ⇒ Calcular el factor de carga UL mediante la fórmula para todos los usuarios a la vez, y posteriormente comprobar si es mayor que 1 para conocer si hay interferencia o no.
- ⇒ Aplicar la fórmula inicialmente para un solo usuario de un servicio, para cada servicio, y así conocer el factor de carga individual de cada servicio. Estableciendo que esos valores obtenidos son el consumo de recursos de cada servicio, se aplica Kaufman-Roberts, conociendo el número de usuarios y teniendo en cuenta que la capacidad total es 1. De esta forma se obtendría la probabilidad de bloqueo a partir de la cual se puede calcular la accesibilidad.

Como ya se ha comentado anteriormente, a nivel cuantitativo no se pueden realizar estos cálculos porque entre las estadísticas del nodo de las que se dispone no se conoce el número de usuarios de cada servicio. Sí se puede calcular cual es el número máximo de usuarios que puede haber de cada servicio (parámetro N_{pole}). Este número es diferente para cada servicio, ya que existe una E_b/N_t

y una tasa de bit diferente en función del tipo de servicio. A partir de estos datos, que se pueden ver en el capítulo anterior, se puede calcular el valor de N_{pole} para cada servicio.

7.8.2. Implementación del algoritmo K-R para el análisis de la potencia del amplificador

Al igual que sucedía con el factor anterior, este tampoco podrá ser calculado por el mismo motivo, es decir, por no disponer del número de usuarios de cada servicio. En este caso se puede calcular la potencia requerida para la transmisión de cada servicio, a partir de las expresiones recogidas en el capítulo anterior. A partir de estos datos se obtiene el consumo de recursos, es decir el consumo de potencia de transmisión (BS_{TXP}), para cada servicio. A continuación se aplica Kaufman-Roberts, considerando el número de usuarios y que la capacidad total es de 40 W, para conocer la accesibilidad.

7.9. Procesado de estadísticas

El esquema de la figura 7.21 muestra el proceso completo del procesado de estadísticas. Los elementos de la red, principalmente nodos B y RNC, reportan periódicamente la información acumulada en sus contadores a la base de datos. Esta información hace referencia a multitud de aspectos, desde el número de celdas ATM enviadas a través de una interfaz hasta la carga de un procesador o el número de conexiones RRC falladas.

Los contadores brutos se transfieren a la base de datos cada 15 minutos; esto es un hecho muy importante y que tiene un impacto crítico sobre el modelo. 15 minutos es un tiempo lo suficientemente largo como para que las variaciones de tráfico sean lo suficientemente importantes como para tener que tratar de buscar una solución que modele el tráfico durante ese tiempo. La importancia radica en que los momentos de congestión se darán en muchos casos debidos a picos de tráfico de escasos minutos, por ejemplo, tras finalizar las campanadas de fin de año, durante el descanso de un partido de fútbol, al finalizar un concierto de música, etc.

Los contadores brutos ofrecen información que en la mayoría de las ocasiones no aporta información relevante al operador de mantenimiento de la red. Es necesario realizar un procesado de los mismos para originar KPIs que aporten información concreta, por ejemplo, a partir de contadores brutos de intentos de llamadas caídas por pérdida de *radio link*, caídas por falta de vecinas, caídas por falta de sincronismo, etc. Se puede formar el KPI tasa llamadas caídas que, aglutinando todas las causas de llamadas caídas y teniendo en cuenta el número de llamadas, diga cuál es la tasa de caídas en un determinado nodo.

Una vez procesadas, los operadores e ingenieros de optimización de red podrán acceder y visualizar la evolución de los diferentes KPIs mediante el uso de herramientas software diseñadas para tal fin. Los KPIs proporcionados por Ericsson han sido tratados con el programa MyCom NIMS-PrOptima.

7.10. KPIs

Se ha definido el término KPI (*Key Performance Indicator*) como una medida de nivel de desempeño de la red en un determinado aspecto. Son datos estadísticos reportados por los nodos

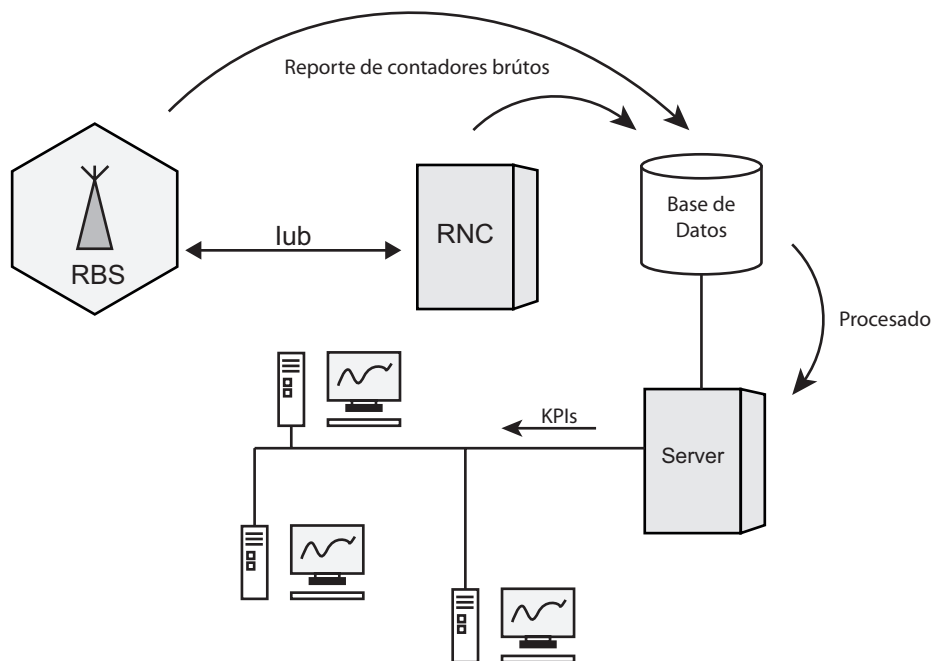


Figura 7.21: Arquitectura del procesado de estadísticas

que, como ya se ha comentado, la mayoría de los fabricantes implementan como datos agregados de cada cuarto de hora (96 muestras diarias).

A continuación se presentan de forma general algunos de los parámetros más importantes. En el apéndice F se pueden encontrar los descriptores específicos de equipos Ericsson junto con respectivos ejemplos reales de su evolución a lo largo de un día.

Estadísticos de tráfico entrante

- ⇒ Peticiones de RAB para circuitos.
- ⇒ Peticiones de cualquier tipo de RAB de paquetes R99.
- ⇒ Peticiones RAB de HSDPA.
- ⇒ Peticiones RAB de HSUPA.
- ⇒ Peticiones rechazadas o falladas de cualquier tipo de RAB de circuitos.
- ⇒ Peticiones rechazadas o falladas de cualquier tipo de RAB de paquetes R99.
- ⇒ Peticiones rechazadas o falladas RAB de HSDPA.
- ⇒ Peticiones rechazadas o falladas RAB de HSUPA.
- ⇒ Minutos voz cursados durante un cuarto de hora.

Estadísticos relacionados con el tráfico de datos

- ⇒ Tráfico en Kilobytes de paquetes HSDPA descargado por todos los usuarios.
- ⇒ Tráfico en Kilobytes de paquetes HSUPA subidos por todos los usuarios.

- ⇒ Tráfico en Kilobytes de paquetes R99 descargado por todos los usuarios.
- ⇒ Tráfico en Kilobytes de paquetes R99 subido por todos los usuarios.
- ⇒ Porcentaje de tráfico cursado en softthandover por unidad de tiempo.
- ⇒ Clientes HSDPA activos por unidad de tiempo.
- ⇒ Clientes HSUPA activos por unidad de tiempo.
- ⇒ Velocidad promedio HSDPA de todas las conexiones realizadas en el nodo.

Estadísticos relacionados con el bloqueo

- ⇒ Accesibilidad total (complementario del bloqueo) de las conexiones sobre circuitos conmutados.
- ⇒ Accesibilidad total (complementario del bloqueo) de las conexiones sobre paquetes conmutados.

7.11. Elección Entorno de Implementación del Modelado

Una vez explicada la estrategia de modelado a seguir, se hace necesario elegir una plataforma sobre la que llevar a cabo la implementación final. A continuación, se describen las opciones barajadas.

7.11.1. Opnet Modeler

La primera opción para llevar a cabo el modelado fue emplear el entorno de simulación Opnet. La razón principal es que permite imitar el comportamiento de un sistema real conforme evoluciona el tiempo.

Opnet proporciona uno de los mejores entornos de simulación orientados a las comunicaciones. Proporciona acceso directo al código fuente, siendo esto una gran ventaja al permitir modificar el comportamiento de las librerías que utiliza el propio software.

Todos los elementos que intervienen en la simulación definen sus características de funcionamiento de acuerdo a una jerarquía de diseño en capas, figura 7.22. En la capa superior se define la topología de la red y los elementos que la compone. Seguidamente, disponemos de un modelo de nodos donde se define la estructura interna de éstos y, por último, tenemos el modelo de procesos donde se definen los estados que definen el nodo.

Para una correcta simulación es necesario definir correctamente las tres capas anteriores, puesto que de lo contrario la simulación sería errónea.

El simulador Opnet se vende de forma modular. Cada uno de ellos constituye una tecnología diferente; así, por ejemplo, se podría adquirir un módulo ZigBee, WiMaX, LTE, etc. En este caso, se necesita el módulo que permite trabajar con redes 3G UMTS, que soporta las siguientes características:

- ⇒ Soporte para 4 clases QoS: *Background*, *Conversational*, *Interactive*, *Streaming*.
- ⇒ Soporte para UE, Repetidor, Nodo B, RNC, SGSN, GGSN con ATM y conexión a redes IP.

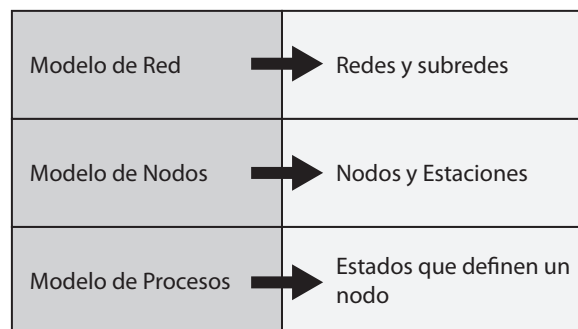


Figura 7.22: Jerarquía de diseño Opnet

⇒ *Dedicated* (DCH) y *Common / Shared Channels* (RACH, FACH, DSCH).

⇒ Multiplexado de canales lógicos en canales de transporte.

⇒ *Acknowledged / Unacknowledged / Transparent RLC modes*.

⇒ *Radio Access Bearer setup, release, negotiation, renegotiation, preemption*.

⇒ Control de admisión *Open Source*.

⇒ *Outer loop power control*.

⇒ *Hard / Soft / Softer handover*.

⇒ GTP soportado hasta la RNC.

El punto fuerte para apostar por Opnet es el potencial de la herramienta, que permite tener en cuenta prácticamente todos los elementos implicados en la red UMTS: se simularía una red completa desde el equipo móvil hasta la GMSC o GGSN. Se podría incluir en el modelo bloqueo por causas CE o falta de códigos, o incluso por problemas hardware ya que podrían simularse estadísticamente.

A su favor también juega la gran capacidad de la que dispone Opnet para precisamente validar modelos: se puede importar trazas de tráfico, crear patrones propios o utilizar algunos proporcionados por el propio programa.

Sin embargo, también existen puntos desfavorables. Opnet implementa la Release 5 de UMTS, por tanto, no incorpora HSUPA que fue introducido en la Release 6. Además, no se proporciona soporte para MSC y GMSC, con lo cual no se tiene el plano de circuitos conmutados, tan sólo está contemplada la parte de paquetes con la excepción de HSUPA.

Debido a que Opnet es una plataforma bastante abierta, sus librerías son accesibles, se podrían ampliar las funcionalidades no incorporadas y contemplar en el simulador la posibilidad de cursar HSUPA y, por supuesto, conexiones de circuitos. Sin embargo, para conseguir este objetivo sería necesario programar cada una de las tres capas de la jerarquía definida por Opnet, desde el nivel más bajo, es decir, a nivel de canales de tráfico UMTS pasando por cómo se tratan las conexiones en los diferentes nodos y, por último, como se mapean los servicios sobre ellas. Sería necesario pues definir por completo los UMTS para estudiar tan solo una parte muy concreta como es el Iub.

Se realizó una primera aproximación al entorno Opnet, figura 7.23, en la que se reprodujo el funcionamiento de una red UMTS y se configuró a fin de obtener resultados fiables. El modelo inicial se componía de un nodo B, la RNC y el SGSN y GGSN que se conectaban a una red local

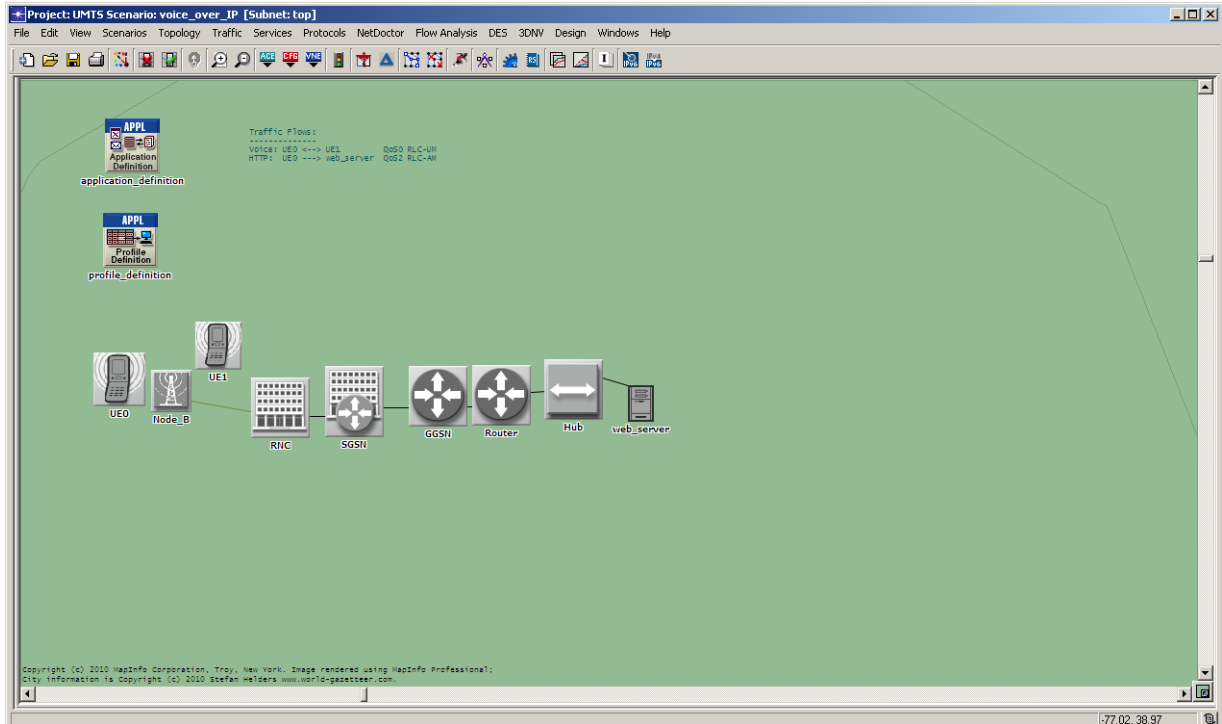


Figura 7.23: Proyecto sobre Opnet

con un servidor al que los usuarios hacían peticiones. La transmisión utilizada fue ATM sobre PDH.

Los resultados no fueron muy satisfactorios ya que apenas se podía controlar el funcionamiento interno del nodo B y la RNC sin tener que reprogramar casi toda su lógica interna, lo que, unido a la carencia de tráfico HSUPA y la falta de tráfico sobre circuitos conmutados, llevó a descartar el uso de Opnet para la Tesis.

7.11.2. Visual Basic for Applications

La siguiente opción barajada fue el empleo de *Visual Basic for Applications*, el lenguaje de macros de Microsoft Visual Basic que se utiliza para programar aplicaciones Windows y que se incluye en varias aplicaciones Microsoft. VBA permite a usuarios y programadores ampliar la funcionalidad de programas de Microsoft Office. Visual Basic para Aplicaciones es un subconjunto casi completo de Visual Basic 5.0 y 6.0.

Microsoft VBA viene integrado en aplicaciones de Microsoft Office, como Word, Excel, Access, Powerpoint y Visio. Prácticamente cualquier cosa que se pueda programar en Visual Basic 5.0 o 6.0 se puede hacer también dentro de un documento de Office, con la sola limitación de que el producto final no se puede compilar separadamente del documento, hoja o base de datos en que fue creado, es decir, se convierte en una macro. Ésta puede instalarse o distribuirse con sólo copiar el documento, presentación o base de datos.

El motivo de utilizar VBA radica en el hecho de que es una herramienta muy extendida en el mundo de la empresa, sobre todo de la mano de la hoja de cálculo Excel. El modelo obtenido podría ser utilizado en el día a día de los ingenieros de transmisión a la hora de dimensionar nuevos enlaces ya que no requeriría de la instalación de programas o módulos adicionales y se ajustaría a

la perfección con los hábitos de trabajo actuales.

Sin embargo, VBA deja poco margen a la hora de tratar algoritmos matemáticos con cierta complejidad, ya que su utilidad principal es automatizar tareas cotidianas, así como crear aplicaciones y servicios de bases de datos para el escritorio. En ningún caso, fue concebido para sustituir a los lenguajes de programación habituales. Este motivo conduce a tratar de buscar otra alternativa que proporcione mayor potencia de procesado.

7.11.3. Programación sobre C# /Java/Python

Si el principal problema que presentaba VBA era su pobre rendimiento, lo óptimo sería pensar entonces en lenguajes de programación de alto nivel que no tengan esas carencias. Entre los lenguajes más populares hoy en día se pueden encontrar Java, C# o Python.

7.11.3.1. C#

C# es un lenguaje de programación orientado a objetos desarrollado y estandarizado por Microsoft como parte de su plataforma .NET, que después fue aprobado como un estándar por la ECMA e ISO.

Su sintaxis básica deriva de C/C++ y utiliza el modelo de objetos de la plataforma .NET, similar al de Java aunque incluye mejoras derivadas de otros lenguajes (entre ellos Delphi).

Con C# se podría crear un programa para la plataforma Windows mediante interfaz nativa de una forma relativamente fácil y limpia gracias al entorno Visual Studio.

7.11.3.2. Java

Java es un lenguaje de programación orientado a objetos, desarrollado por Sun Microsystems a principios de los años 90. El lenguaje en sí mismo toma mucha de su sintaxis de C y C++, pero tiene un modelo de objetos más simple y elimina herramientas de bajo nivel, que suelen inducir a muchos errores, como la manipulación directa de punteros o memoria.

Las aplicaciones Java están típicamente compiladas en un bytecode, aunque la compilación en código máquina nativo también es posible. En el tiempo de ejecución, el bytecode es normalmente interpretado o compilado a código nativo para la ejecución, aunque la ejecución directa por hardware del bytecode por un procesador Java también es posible.

La gran ventaja de emplear Java es la capacidad de adaptación al entorno: puede asegurarse que el programa funcionará en prácticamente cualquier dispositivo con máquina virtual java.

7.11.3.3. Python

Python es un lenguaje de programación de alto nivel cuya filosofía hace hincapié en una sintaxis muy limpia y que favorezca un código legible.

Se trata de un lenguaje de programación multiparadigma ya que soporta orientación a objetos, programación imperativa y, en menor medida, programación funcional. Es un lenguaje interpretado, usa tipado dinámico, es fuertemente tipado y multiplataforma.

Python ofrece generalmente la posibilidad de generar programas con gran legibilidad y transparencia, permite recorrer el código de manera intuitiva y didáctica, lo que podría ser un punto favorable a la hora de abordar nuevos trabajos en futuros proyectos.

Emplear cualquiera de estos lenguajes, u otro, implica centrar esfuerzos en aspectos que se alejan totalmente del objetivo: estudiar el canal Iub. La programación de interfaces gráficas que permitan mostrar resultados, la programación de algoritmos matemáticos, el tratamiento de la información, etc., son aspectos que no aportan nada al estudio que se está realizando. Es necesaria una herramienta que reúna la capacidad de hacer una programación lo suficientemente potente sin que ésta entorpezca el objetivo, a la vez que ofrece capacidad para presentar los resultados de una manera sencilla.

7.11.4. Mathworks MATLAB

MATLAB es un software matemático que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio (lenguaje M). Entre sus prestaciones básicas se hallan: la manipulación de matrices, la representación de datos y funciones, la implementación de algoritmos, la creación de interfaces de usuario (GUI) y la comunicación con programas en otros lenguajes y con otros dispositivos hardware.

Con MATLAB estará disponible un lenguaje de programación lo suficientemente potente, a la vez que sencillo, para implementar los algoritmos. Además dado que se emplean diversas funciones matemáticas, y MATLAB es fundamentalmente un software matemático, no se desperdiciará tiempo ni esfuerzo en desarrollar funciones que ya vienen integradas en el propio programa. Junto con el lenguaje M, MATLAB ofrece muy buenas herramientas para generar gráficas que puedan ser usadas para contrastar los datos reales procedentes de estadísticas.

Así pues, y dado el carácter de investigación que tiene la Tesis, MATLAB será el escenario elegido para implementar el programa.

7.12. Conclusiones

Una vez descrito el proceso de modelado general Kaufman Roberts aplicado al consumo de recursos en redes multiservicio, en este capítulo se particulariza para los recursos que serán susceptibles de ampliación durante la operación de la red. Así, será posible determinar el consumo de capacidad en las tarjetas de procesado banda base (en términos de Channel Elements). Así, se sabrá si un nodo tiene congestión por falta capacidad de procesado en el Uplink o en el Downlink, o bien si se tienen rechazos debidos a un fallo hardware en estos elementos. Será posible incluso ampliarlos de forma proactiva una vez conocida la previsión de tráfico para determinados eventos en la red. En cuanto al consumo de códigos de spreading en el Downlink, se podrá determinar si es necesaria capacidad adicional para evitar bloqueo. En caso afirmativo, se deberá añadir una nueva portadora WCDMA (cada una de ellas contendrá un árbol completo de códigos).

Como se verá posteriormente, para la presente Tesis se han definido una serie de situaciones o escenarios que servirán para poner a prueba el modelo teórico. La diferencia entre todos ellos radica fundamentalmente en el tráfico y los recursos de que se dispone. Así será posible testear el comportamiento del sistema frente a congestión bajo diferentes supuestos.

En cada uno de esos escenarios se han realizado ensayos independientes en función del modelado de tráfico según las distribuciones comentadas anteriormente. Al mismo tiempo, los parámetros implícitos en cada una de ellas (media, varianza, λ ...) se ajustan de acuerdo a métricas distintas entre la accesibilidad real y la estimada:

⇒ Minimizando el error cuadrático medio (RMS).

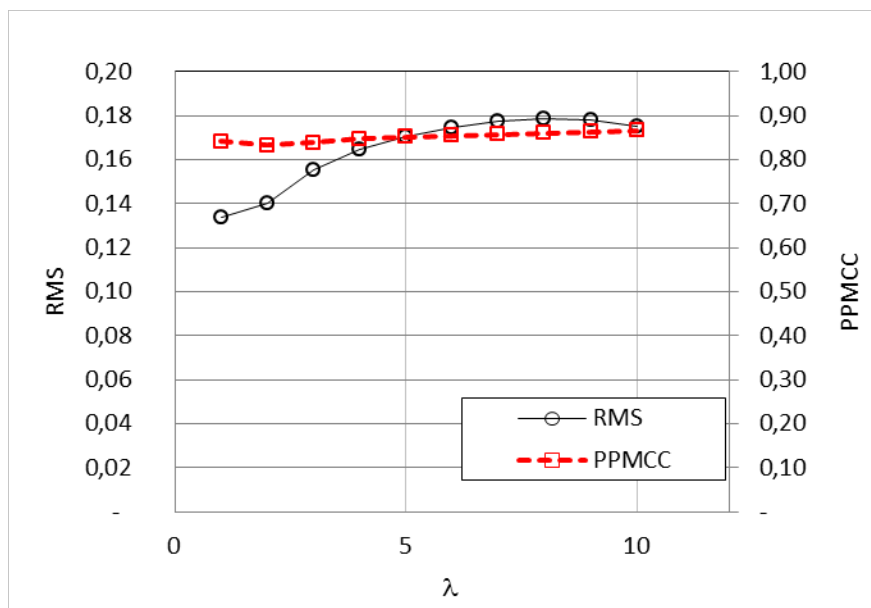


Figura 7.24: Evolución de las métricas RMS y PPMCC para el caso de tráfico interpolado según distribución de Poisson.

⇒ Optimizando el coeficiente de correlación de Pearson (PPMCC).

⇒ % de error en la reconstrucción de la muestra original.

Cuando el tráfico se modela según una distribución de Poisson se hace variar su parámetro λ entre 1 y 10. En la figura 7.24 puede verse la evolución de las 2 métricas para los diferentes valores de λ en un escenario. Como se ha comentado anteriormente, el hecho de eventanar la curva de densidad de probabilidad provoca que tras la reconstrucción de la muestra original se cometa un cierto error.

Igualmente, cuando en ese escenario el tráfico se modela según una distribución normal existen dos parámetros variables, su media y su varianza. Para simplificar el estudio, se fija un valor de media $\mu = 7,5$, valor intermedio dentro del intervalo $[0, 15]$, y se hacen diferentes simulaciones variando el valor de su varianza. La evolución de las métricas según este procedimiento puede verse en la figura 7.25.

Como puede observarse, el error cuadrático medio y el coeficiente de correlación de Pearson presentan comportamientos muy similares en cualquiera de las dos distribuciones. A la vista de los resultados obtenidos, se elige como curva para la interpolación de tráfico la distribución de Poisson con parámetro $\lambda = 1$. De esta manera, el problema se ajusta al modelado de un sistema de comunicaciones típico en el que la tasa de llegadas sigue una distribución exponencial y los clientes llegan según un proceso de Poisson. Además, eligiendo esta distribución se elimina uno de los dos grados de libertad, ya que la media y la varianza son iguales a λ .

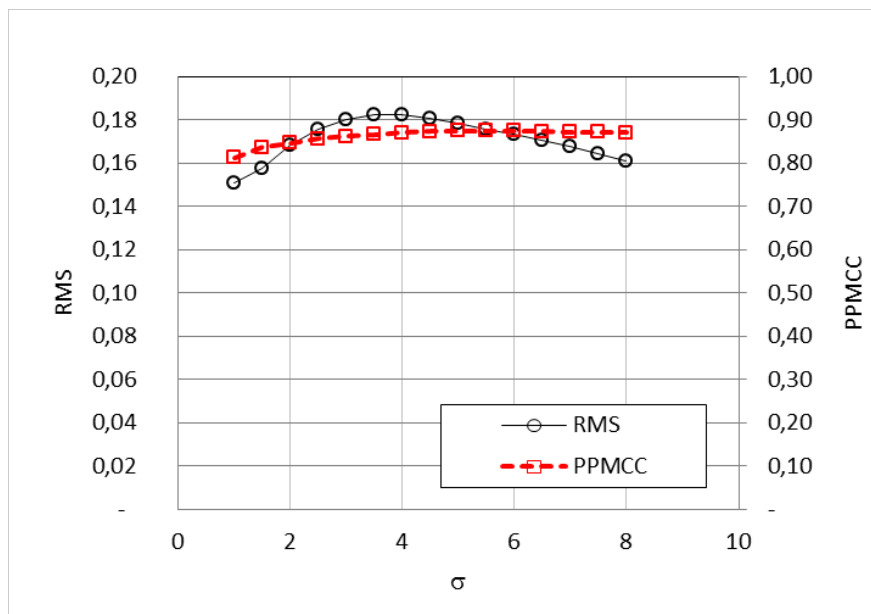


Figura 7.25: Evolución de las métricas RMS y PPMCC para el caso de tráfico interpolado según distribución normal.

Capítulo 8

Comparativa entre los sistemas UTRAN modelado y real

8.1. Accesibilidad binaria y medida de calidad

8.1.1. Definición de calibración

A partir de las observaciones realizadas en el capítulo anterior, la calibración del sistema tiene por objeto obtener un modelo teórico que reproduzca de la forma más fiel posible el comportamiento real. Dicho de otro modo, conocidos los recursos y el tráfico, que el modelo sea capaz de dictaminar si una degradación es debida a congestión o a fallos hardware.

De forma ideal, el modelo estaría perfectamente calibrado si reprodujese al 100 % el comportamiento real. Como resta decir que la realidad en raras ocasiones se ajusta a modelos matemáticos rígidos, es perfectamente asumible que se producirá una desviación más o menos significativa entre los resultados teóricos y las observaciones reales.

El enfoque que se plantea en este momento es definir un marco de calidad y verificar si el modelo teórico se ajusta a dicho contexto. En caso afirmativo, el modelo se dirá que está calibrado; en caso negativo, se buscará realimentar el sistema modificando alguno de sus parámetros, hasta conseguir que éste alcance los mínimos de calidad.

8.1.2. Enfoque lógico o binario

En primera instancia, más que reproducir fielmente unas cantidades numéricas de accesibilidad real vs. accesibilidad teórica, se trata de que el modelo teórico detecte la degradación al mismo tiempo que esta se produce. Por tanto, se puede llevar el problema al terreno de la lógica binaria de la siguiente forma: «0» significa que no hay degradación y «1» que sí la hay. La tabla 8.1 recoge pues las 4 situaciones que podrían darse durante la calibración del sistema.

Estos 4 casos, combinación de las variables binarias anteriormente definidas, se pueden recoger en dos:

- ⇒ Éxito: se define como una situación en la que la predicción teórica coincide con la situación real. Se da cuando existe degradación real y el modelo así la predice, o cuando no existe degradación y el modelo así lo determina.

$Acc_{teórica}$	Acc_{real}	Caso
SI	SI	Buen funcionamiento
SI	NO	Falsa alarma
NO	SI	No detectado
NO	NO	Buen funcionamiento

Cuadro 8.1: Situaciones posibles durante la calibración del sistema

	A_r
$Acc_{real} < \eta_r$	1
$Acc_{real} \geq \eta_r$	0

Cuadro 8.2: Definición de la variable lógica A_r

\Rightarrow Discrepancia: se define como una situación en la que la predicción teórica no coincide con la situación real. Existen dos posibilidades: que habiendo degradación, el modelo teórico no la detecte, y el caso de las falsas alarmas cuando el modelo predice una degradación que en realidad no existe. La calibración del sistema pasa por minimizar la tasa de discrepancias, de manera que se dirá que la calibración es de calidad cuando la tasa de errores sea menor de un cierto valor (por ejemplo, del 6%).

En la práctica, los ingenieros de red de las distintas operadoras de UMTS consideran que existe una anomalía en el nodo cuando el reporte de accesibilidad desciende por debajo de un umbral $\eta_r < 99\%$. Por tanto, puede definirse una variable lógica $A_r = \{0, 1\}$ según la tabla 8.2.

Por otra parte, el parámetro de accesibilidad ligado a la accesibilidad teórica calculada se puede discretizar definiendo una variable lógica $A_t = \{0, 1\}$ bajo el mismo enfoque: se detectará una anomalía cuando ese valor descienda por debajo de un determinado valor umbral η_t . Este umbral de decisión puede particularizarse en función del escenario, ya que independientemente de los valores que vaya tomando la accesibilidad, a la vista de su evolución el ojo humano puede determinar cuándo se está produciendo efectivamente la degradación. En la figura 8.1 se ilustra esta idea: de forma intencionada, el autor no ha rotulado numéricamente el eje de abscisas en la gráfica: los números pueden diferir de un contexto a otro, pero el momento en que se produce la degradación está perfectamente definido. Por tanto, a priori cada curva de accesibilidad puede tener su propio nivel o umbral de decisión.

8.1.3. Medida de calidad

Una vez definidas las variables lógicas con las que se trabajará en el estudio, es necesario definir de antemano una política de calidad. De esta forma se evalúa el modelo teórico en contraposición con la realidad y, como se verá posteriormente, se modificarán sus variables hasta conseguir que cumpla con los objetivos de calidad.

Se utilizará como medida de calidad la tasa de discrepancias entre las accesibilidades lógicas teórica A_t y real A_r . La manera de medir cuándo se producen discrepancias entre el modelo y la

	A_t
$Acc_{teórica} < \eta_t$	1
$Acc_{teórica} \geq \eta_t$	0

Cuadro 8.3: Definición de la variable lógica A_r

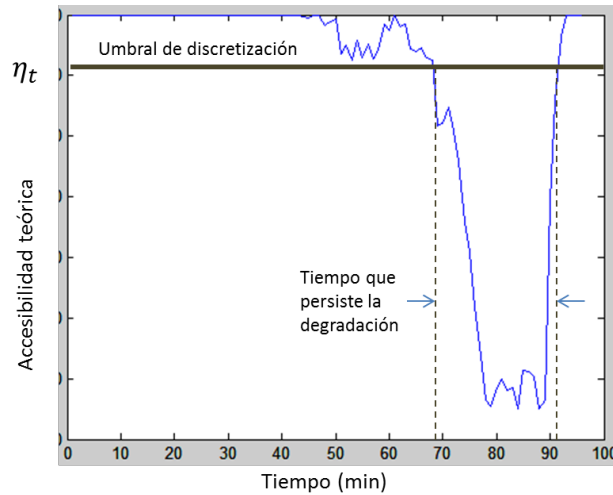


Figura 8.1: Umbral teórico de decisión de una accesibilidad teórica

A_t	A_r	$\delta_i = XOR(A_t, A_r)$	Caso
0	0	0	Acierto (situación de no degradación)
0	1	1	Discrepancia (fallo de detección)
1	0	1	Discrepancia (falsa alarma)
1	1	0	Acierto (degradación detectada)

Cuadro 8.4: Operación XOR para la detección de discrepancias

realidad es utilizando alguna función lógica que permita distinguir cuándo dos variables lógicas son diferentes. Una posibilidad es utilizar la función lógica OR-Exclusiva, o XOR. En la tabla 8.4 se recogen las distintas combinaciones δ_i que pueden obtenerse, de manera que cuando $\delta_i = 0$ el modelo ha acertado, mientras que si $\delta_i = 1$ se ha producido una discrepancia. Una vez determinado muestra a muestra dónde se producen las discrepancias, la ecuación 8.1 recoge la manera de calcular una tasa de error Δ en un vector δ de longitud 96, es decir, los datos obtenidos para todo un día.

$$\Delta (\%) = \frac{\sum_{i=0}^{96} \delta_i}{96} \cdot 100 \quad (8.1)$$

Se plantea como objetivo de calidad obtener un modelo teórico con una tasa de discrepancias $\Delta < n\%$. Dado que el valor lógico de A_t depende del umbral de decisión η_t se tratará de adaptar el valor de éste para minimizar la tasa Δ y lograr así un modelo con la calidad y fiabilidad requeridas.

Por último, en la figura 8.2 se ilustra un ejemplo que sintetiza todas las operaciones vistas anteriormente. En la gráfica (a) se representa un ejemplo de accesibilidad teórica, y la operación de decisión según un umbral η_t del 70%, que da como resultado una señal digital A_t . Como se puede apreciar, todas aquellas muestras cuyo valor de accesibilidad cae por debajo del 70% se convierten en un «0» lógico, mientras que al resto les corresponde «1». Por otra parte, en la gráfica (b) se repite el mismo proceso para obtener una señal digital A_r pero digitalizando la accesibilidad real reportada por el nodo. En este caso el umbral de decisión η_r suele estar prefijado por los operadores en un 99% (aunque el proceso sería el mismo si este número fuera cualquier otro). La correspondencia lógica binaria es la misma en los dos casos. Por último, en la gráfica inferior se ilustra el proceso de búsqueda de discrepancias, realizando la operación lógica XOR entre las dos señales digitales previas, A_t y A_r . La señal digital que se obtiene, Δ , refleja con «1» aquellos

instantes en los que existen discrepancias, y con «0» las coincidencias. Por tanto, sumando los elementos del vector Δ se tiene el número exacto de discrepancias.

En resumen, el proceso explicado anteriormente, representado en la figura 8.3 se compone de las siguientes etapas:

1. A partir de los recursos y del tráfico reportado por el Nodo B, se calculan las accesibilidades correspondientes a los enlaces downlink y uplink (bloques con líneas punteadas). Implícitamente, este proceso se compone de 3 subprocesos:
 - a) Interpolación del tráfico reportado por el Nodo B.
 - b) Aplicación en paralelo del algoritmo de Kaufman-Roberts a 3 recursos (Iub, CE, SF).
 - c) Multiplicación de las 3 accesibilidades parciales derivadas del punto anterior.
2. Del paso anterior se obtienen dos cantidades: accesibilidad en el enlace ascendente (ACC_{DL}) y en el enlace descendente (ACC_{UL}). Ambos parámetros son combinados, tal y como se ha visto anteriormente. Se obtiene pues una accesibilidad teórica, una combinación de las anteriores similar a la que realiza implícitamente el Nodo B al calcular la accesibilidad real.
3. Se produce un digitalizado de la accesibilidad teórica del paso anterior y de la real, con sendos umbrales de decisión η_t y η_r , obteniéndose dos señales digitales A_t y A_r , respectivamente.
4. Las señales digitales anteriores se combinan con la operación lógica «XOR» para detectar las posibles discrepancias, cuya tasa representa el parámetro Δ .

Este proceso permite realizar un primer cálculo sobre la capacidad del modelo de predecir la degradación del sistema. Se trata de un modelo estático, que proporciona una medida de calidad a partir de unos parámetros de entrada como son el tráfico y los recursos disponibles. No obstante, este modelo tiene algunas variables que es necesario concretar:

⇒ Umbral de decisión η_t .

⇒ Parámetro λ de distribución de Poisson para interpolación del tráfico.

Como se verá posteriormente en el apartado 8.3, el umbral de decisión puede determinarse a partir de simulaciones en diferentes escenarios, ya que existe un valor óptimo para todos ellos. Para la determinación del parámetro λ se recurrirá a un sistema de calibración en lazo cerrado, cuyo desarrollo puede encontrar el lector en el apartado 8.3.2.

8.2. Descripción de los casos de uso

A lo largo de la presente Tesis se han considerado diferentes escenarios con tráficos, recursos y configuraciones diferentes. Los casos de uso se han elegido de tal manera que reflejen situaciones límite en las que los servicios de UMTS experimenten bloqueo, o lo que es lo mismo, se haya determinado que el nodo se encuentra en estado degradado. En UMTS se denomina nodo degradado a los nodos que presenten un bloqueo superior al 1 %, o lo que es lo mismo, una accesibilidad del 99 %. Los casos de estudio con recursos limitados son extendidos con un caso complementario que incluya el mismo escenario, en otro instante de condiciones ambientales similares, pero tras haber realizado actualizaciones al nodo, aliviando el recurso que se creía saturado. De esta forma, se dispone de los datos reales suficientes para:

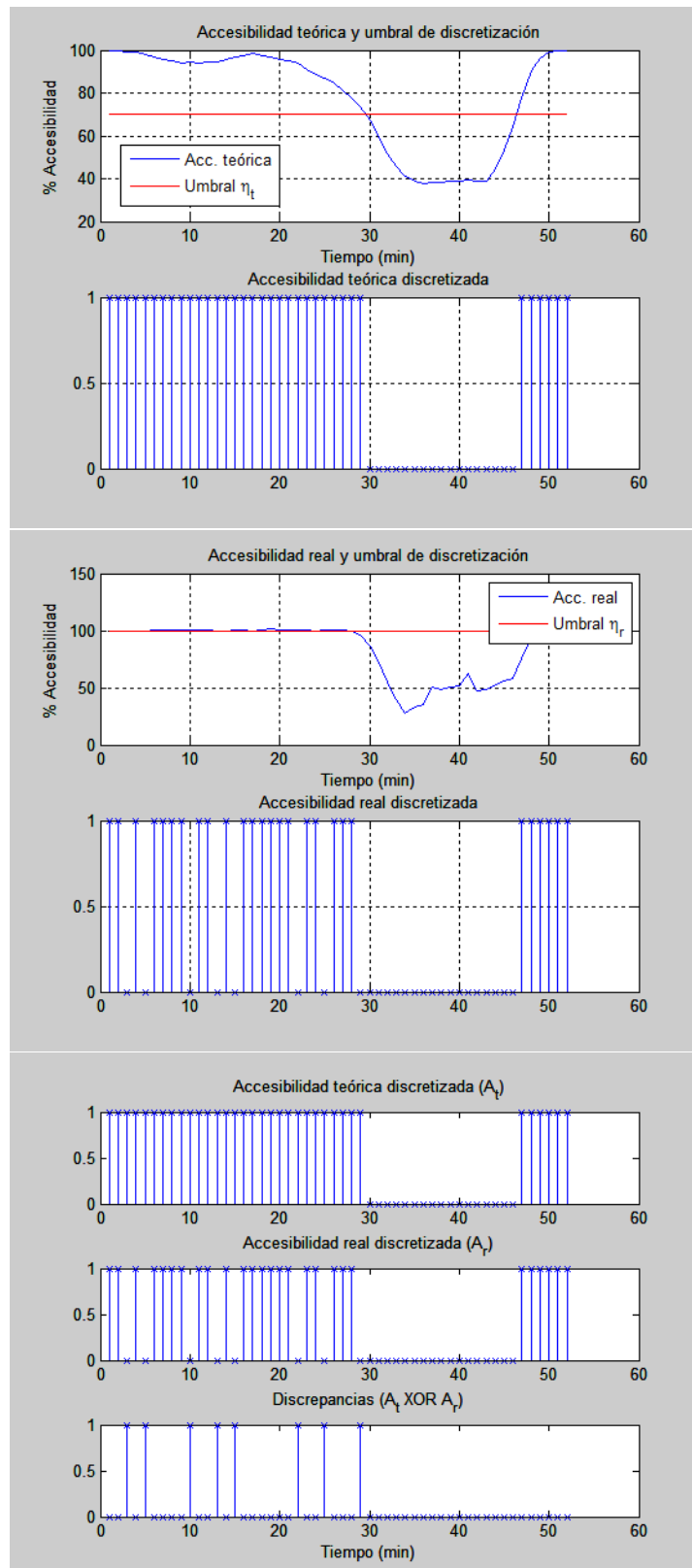


Figura 8.2: Ejemplo de discretización y búsqueda de discrepancias

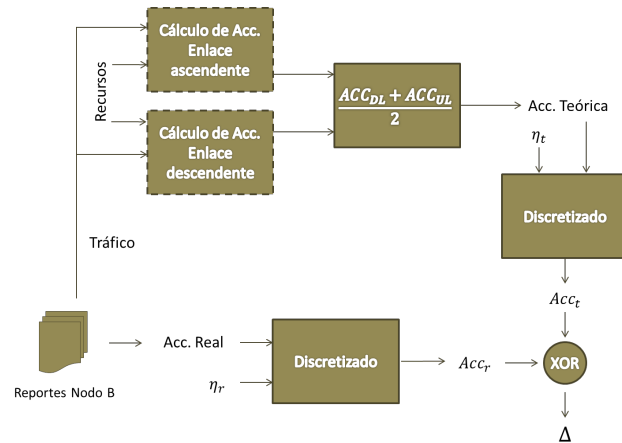


Figura 8.3: Esquema de comparación accesibilidades teórica y real

⇒ Calibrar el tráfico.

⇒ Probar la eficacia de las predicciones del modelo.

Cada caso de uso se compone de dos escenarios. Los escenarios se corresponden a las situaciones independientes del mismo caso de uso, para condiciones de entorno equiparables. Se consideran condiciones equiparables la utilización del mismo emplazamiento en el mismo día laboral de dos semanas sucesivas en el tiempo. El primero de los escenarios se corresponde con la situación en la que el nodo bajo estudio presenta alguna limitación que ocasiona el bloqueo de algún servicio. El segundo de los escenarios se corresponde con los datos recogidos en el mismo nodo en las condiciones descritas, una vez que se ha realizado una actuación sobre el nodo, suponiendo ésta una mejora de la accesibilidad resultante.

A lo largo de este documento se ha hecho referencia en numerosas ocasiones a la existencia de 3 recursos limitantes básicos, Iub, Channel Elements y códigos de ensanchamiento (SF), a los que hay que añadir factores adicionales que también limitan la capacidad del sistema UMTS. El modelo desarrollado obtiene la accesibilidad estimada teniendo en cuenta todos los factores limitantes que se han modelado.

Los casos de uso analizados buscan evaluar las predicciones del modelo diseñado en situaciones conocidas. Éstos se eligen para evaluar las diferencias en la accesibilidad antes los cambios realizados en los recursos limitantes básicos.

8.2.1. Características generales de los casos de uso

El modelo de estación base es compartido para todos los casos de estudio analizados. No así la configuración del mismo. Entre los factores de configuración comunes a todos los casos de estudio analizados se encuentran:

⇒ Factor de potencia.

⇒ Recogida y procesamiento de KPI.

⇒ Modelos de tarjetas CE (el número podrá variar).

⇒ Posibilidad de configuración Dual Stack o Dual IP.

⇒ Enlaces ATM E1 (2Mbps).

⇒ Enlaces IP 23Mbps.

Los cálculos de la accesibilidad se hacen siguiendo los principios expuestos en capítulos anteriores. En todos los escenarios se determinará la accesibilidad tanto para los tráficos elásticos (PS) como para los rígidos (CS), tanto para el enlace descendente (DL) como ascendente (UL). Según se ha visto, el cálculo de la accesibilidad total reportada por la red supone la agregación de la accesibilidad en los canales UL y DL. Para el cálculo de la accesibilidad UL se tendrán en cuenta los factores siguientes:

⇒ Iub (PS/CS) según corresponda.

⇒ CE.

Para el cálculo de la accesibilidad DL se tendrán en cuenta los factores siguientes

⇒ Iub (PS/CS) según corresponda.

⇒ CE.

⇒ SC.

8.2.2. Particularidades de los casos de uso

Cada caso de uso tendrá una configuración particular de la estación base, atendiendo a las características específicas de cada emplazamiento particular. En cada caso de uso el nodo estudiado presentaba algún tipo de limitación desconocida que origina una degradación. Cada caso de uso se compone de dos escenarios, el escenario original que demostraba alguna limitación y el escenario ampliado. En el apartado F.4 se detalla la configuración particular de cada escenario en términos de recursos disponibles. En la tabla 8.5 se detalla el dimensionamiento de los recursos para cada uno de los escenarios considerados.

Escenario/Recursos	CE UL	CE DL	SC	Iub UL	Iub DL
1	2x384	1x384 + 1x384	2 portadoras	1 VCC	IP 23 Mbps
2	2x384	1x384 + 1x384	2 portadoras	2 VCC	IP 23 Mbps
3	2x384	1x384 + 1x384	2 portadoras	1 VCC (5 E1)	IP 23 Mbps
4	2x384	1x384 + 1x384	2 portadoras	2 VCC (5 E1)	IP 23 Mbps
5	2x384	2x384	1 portadoras	1 VCC (1 E1)	IP 30 Mbps
6	3x384	2x384	1 portadoras	1 VCC (1 E1)	IP 30 Mbps
7	2x384	2x384	1 portadoras	IP 30 Mbps	IP 10 Mbps
8	3x384	2x384	1 portadoras	IP 30 Mbps	IP 10 Mbps

Cuadro 8.5: Recursos para cada escenario

8.3. Calibración del tráfico

En el capítulo anterior se ha presentado al lector la problemática de trabajar con una resolución temporal baja de acuerdo con los KPIs reportados por el nodo, justificando así el desarrollo de un método de interpolación que permite incrementar la resolución minuto a minuto. También se ha recalcado que la elección de una función de distribución u otra no condiciona tanto el modelo como

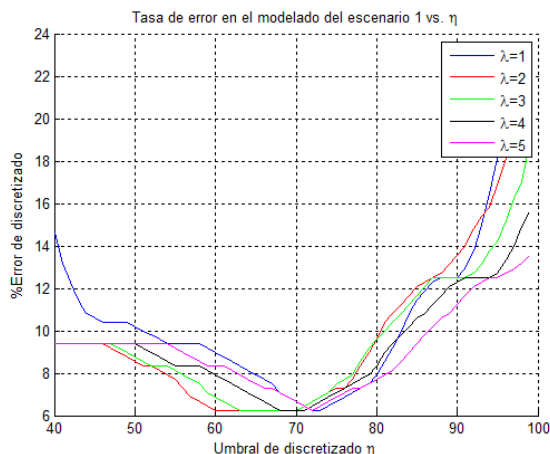


Figura 8.4: Escenario 1, $\Delta = f(\eta_t, \lambda)$

cabría esperar, si bien a efectos prácticos y por coherencia con la bibliografía de teoría de colas y teletráfico se ha decidido utilizar una distribución de Poisson.

Al mismo tiempo, y de cara a construir un modelo efectivo, se ha definido una política de calidad que tiene en cuenta la precisión en la detección de la degradación que pueda arrojar el modelo teórico en contraposición con la situación real. La tasa de discrepancias Δ depende del valor de accesibilidad teórica lógica y, por tanto, de cuál sea el umbral de discretización η_t . En última instancia, la accesibilidad teórica depende cuantitativamente de la distribución de tráfico considerada y, por ello, es función del parámetro λ .

A la vista de las observaciones anteriores, el diseño del modelo tiene como objeto en esta etapa la optimización del binomio (η_t, λ) que minimice la tasa de discrepancias Δ acotada dentro de los parámetros de calidad.

Para la optimización de estos parámetros se recurre a los diferentes casos de uso descritos en el apartado anterior, con objeto de estudiar su dependencia de las condiciones o si, por el contrario, es posible generalizar unos valores concretos.

8.3.1. Elección del umbral de decisión

Se realizan en este punto diferentes simulaciones como combinación de los siguientes parámetros:

- ⇒ Simulaciones para los escenarios 1 a 5.
- ⇒ Para cada escenario, modelos de tráfico con parámetro λ entre 1 y 5.
- ⇒ Para cada caso, umbrales de decisión η_t entre 40 % y 99 %.

Se evalúa para todos los casos la tasa de discrepancias Δ , buscando en todo caso su valor mínimo. En la figura 8.4 se representa, en el escenario 1, la tasa de discrepancias en función del umbral elegido para diferentes valores de λ . Se observa que la tasa de discrepancias alcanza en todos los casos su valor mínimo para valores de umbral en torno al 70 %.

Por otra parte, se estudia la tasa de discrepancias en función de los escenarios. En la figura 8.5 se representa a modo de ejemplo la dependencia de la tasa de errores en los escenarios 1 a 5 para un valor prefijado de distribución de Poisson, por ejemplo $\lambda = 1$. Como sucede en el caso anterior, un umbral de decisión $\eta_t = 70$ % garantiza una tasa de discrepancias Δ mínima.

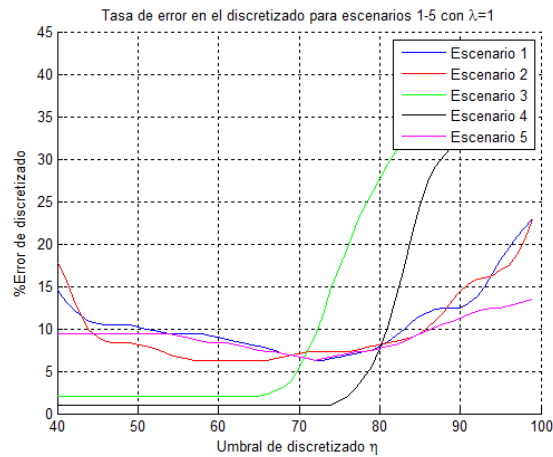


Figura 8.5: Tasa de discrepancias $\Delta = f(\eta_t, \text{escenario})$, $\lambda = 1$

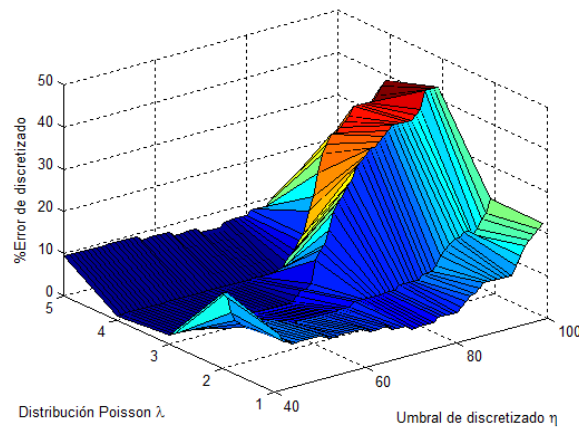


Figura 8.6: Representación tridimensional de $\Delta = f(\eta_t, \lambda)$ en el esc. 1

Por último se complementa con la figura 8.6 una representación tridimensional de la tasa de discrepancias en función del binomio (η_t, λ) para el escenario 1.

Por tanto, a la vista de los resultados se decide que el modelo «digitalice» la accesibilidad teórica según un umbral de decisión $\eta_t = 70\%$ y posteriormente calibrar el modelo variando el parámetro λ , tal como se explicará en apartados posteriores.

8.3.2. Modelado de tráfico como sistema realimentado o en lazo cerrado

8.3.2.1. Introducción a los sistemas realimentados

Como rama de la ingeniería, la regulación automática resulta de gran interés en aquellos procesos industriales en los que es necesario controlar el valor que toma una determinada variable (sistemas de control). Para ello, se parte del modelado de un sistema a partir de subsistemas sencillos entendidos como diagramas de bloques, sobre los cuales se traza un flujo determinado y se añaden otros subsistemas específicos para dotar de control al sistema.

Dentro de la teoría general del análisis dinámico de sistemas, resultan especialmente importantes los sistemas realimentados y, más concretamente, aquellos con realimentación negativa. Este tipo

de modelo permite abstraer el control de una variable teniendo presente un valor de referencia, y produciéndose el control en base a la monitorización del valor de salida del sistema, realizando las operaciones necesarias para garantizar su estabilidad en torno al valor deseado.

El clásico ejemplo es el control del nivel de agua de un depósito. Imagínese el lector que se tiene un depósito con una entrada de agua gobernada por una electroválvula y una salida que puede ser gobernada de forma manual por el usuario o con una electroválvula. Si se desea que el nivel de agua permanezca siempre en un determinado valor, el sistema de control monitorizará la cantidad de agua de que dispone el depósito mediante un sensor. Si el nivel de agua está por debajo de la referencia, el sistema lo detectará y dará la orden de abrir la electroválvula de entrada. Al mismo tiempo, se continúa monitorizando en tiempo real el nivel de agua y cuando se detecta que ya ha alcanzado el nivel de referencia, enviará una señal a la entrada para cerrar la electroválvula. Lo mismo sucede en el caso contrario, si por algún motivo (por ejemplo si el depósito está a la intemperie y se permite la entrada de agua de lluvia) el sistema detecta más agua de la debida, puede abrir la electroválvula de salida automáticamente y seguir monitorizando el nivel de agua hasta que alcance el nivel de referencia.

En la figura 8.7 se representa un modelo general de sistema realimentado. En la práctica alguno de estos bloques podría subdividirse o suprimirse en función de su contexto.

- ⇒ El primer bloque («entrada») representa únicamente la señal que se debe tener como referencia (que previsiblemente proviene de otro subsistema donde ha sido generada pero que en este apartado no será tenida en cuenta).
- ⇒ El último bloque («salida») representa únicamente la señal de salida que se tiene en cada momento, que será o no la deseada. Precisamente de ahí parte hacia abajo un lazo de realimentación que significa que esa señal está siendo monitorizada.
- ⇒ El bloque del lazo («retroalimentación») es el que se utiliza para procesar la señal de salida si fuera necesario, para poder compararla con la señal de referencia.
- ⇒ La unidad de control es el subsistema encargado de comparar la señal de referencia («entrada») con la señal de salida (previo procesado en el bloque de retroalimentación si procede). La comparación que se realiza en este bloque puede ser de muchas maneras. Por ejemplo, para una señal analógica puede efectuarse una resta y comprobar así si dos señales son iguales o si una de ellas es mayor que un nivel de referencia, y para señales digitales se puede utilizar cualquier operación lógica (suma, producto, negación, exclusión...) entre dos señales para comprobar si son coherentes.
- ⇒ Por último, el «actuador» es el subsistema encargado de modificar la señal de salida para tratar de llevarla a los niveles deseados. Este bloque entrega una señal «de salida» que es la que se monitoriza permanentemente para conseguir que la señal de salida sea coherente con la de entrada.

8.3.2.2. Particularización al modelo

Como se ha visto en el apartado anterior, la teoría de los sistemas realimentados permite establecer un modelo de control que permite disponer de un sistema en el cual una variable es monitorizada para que se comporte de acuerdo a unos determinados parámetros.

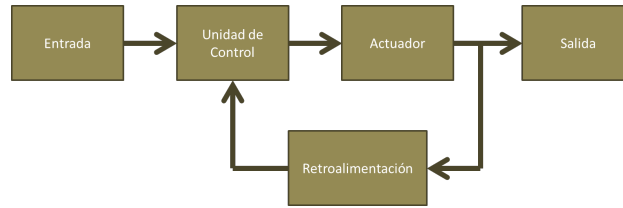


Figura 8.7: Modelo general de un sistema realimentado

Supóngase un vector \vec{v} en el que se almacenan las medidas que se van tomando de una variable A . Si dicha variable tiene una dependencia temporal, $A = A(t)$, entonces la posición i -ésima del vector v_i viene definido por la expresión 8.2:

$$v_i = A(t), t = i \quad (8.2)$$

Al margen de una implementación real del vector anterior, en general los sistemas realimentados como el ejemplo del depósito explicado en el apartado anterior, se adaptan a la expresión anterior. En este caso, cada una de las medidas no depende de diferentes instantes de tiempo, sino que la variable a controlar irá tomando diferentes valores en función del escenario, de acuerdo con la expresión 8.3. Así, el elemento i -ésimo del vector, v_i representará el valor de la variable en el escenario i -ésimo.

$$v_i = A(esc), esc = i \quad (8.3)$$

La realimentación será necesaria por tanto a medida que se van introduciendo nuevos escenarios en el estudio, esperando encontrar un patrón de tráfico estable que permita al modelo alcanzar su grado de calidad (recuérdese que la interpolación del tráfico con una distribución de Poisson de parámetro λ pertenece al modelo).

Téngase en cuenta que si se pretende que el sistema haga una predicción de la degradación con el mínimo error posible, es imprescindible que la tasa de discrepancias entre el modelo teórico y el real esté por debajo de un determinado valor (que en el capítulo anterior se ha definido como criterio de calidad). El objeto en este apartado es por tanto construir un modelo realimentado que permita calibrar el subsistema dedicado a la interpolación del tráfico. En la figura 8.8 se representa un sistema en lazo cerrado particularizado para el calibrado del tráfico. En apartados posteriores se detalla su implementación.

Señal de entrada

La señal de entrada se corresponde con aquella referencia que se desea tomar. En este caso, se desea que la calidad del sistema permanezca estabilizada tanto como sea posible, de manera que como señal de entrada se tomará un valor de tasa de discrepancias $\Delta_0 = 6\%$.

Señal de salida

Esta señal es aquella que se monitoriza permanentemente con objeto de estabilizar el sistema. Generalmente es un valor de la misma naturaleza, homólogo a la señal de entrada. En este caso, se observará cuál es la tasa de discrepancias real que se obtiene con el modelo actual, Δ .

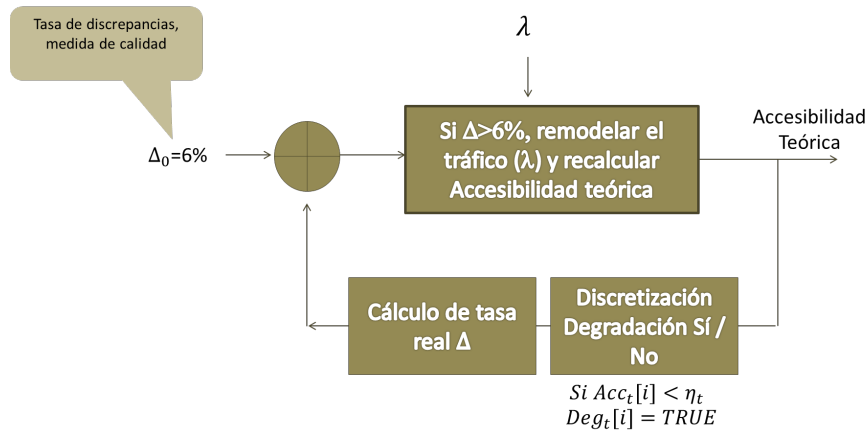


Figura 8.8: Particularización de modelo realimentado para la calibración del tráfico

Retroalimentación

Este bloque a priori tiene como objeto asegurar la coherencia entre la señal de salida y la señal de referencia, para que de esta forma sean comparables. Si a la salida del actuador se tiene la accesibilidad teórica, en este bloque puede implementarse la digitalización de la señal A_t y el cálculo de la tasa de discrepancias Δ .

Unidad de control

Bajo este subsistema se encuentran todas aquellas operaciones que es necesario realizar para determinar si el sistema está estabilizado en torno a las condiciones impuestas. Compara la señal de referencia con la salida (adaptada si es necesario) y en función del resultado de la comparación se desencadenan unas determinadas acciones. En este caso se trata de comparar el binomio (Δ, Δ_0) , comprobando así si la tasa de discrepancias es suficientemente baja; en definitiva, verificar si el sistema es de calidad.

Actuador

Englobará todas aquellas acciones que es necesario tomar para modificar la señal de salida, una vez que se ha determinado que el sistema no tiene la calidad suficiente. En definitiva la misión del actuador es modificar las variables del sistema. Teniendo en cuenta que, como se ha visto anteriormente, el umbral de decisión η_t no varía significativamente con la distribución ni con el escenario, en la calibración del modelo se modificarán los parámetros de interpolación del tráfico, λ . Corresponde a la implementación siguiente.

8.3.2.3. Detalles de implementación

A nivel macroscópico, el contexto en el que se ubica el sistema de calibración implementado puede verse esquematizado en la figura 8.9.

Como ya se ha comentado anteriormente, un valor de λ determinado da como resultado un modelo de calidad para un escenario concreto. Evidentemente, el modelo será mucho más efectivo si se consigue enunciar de un modo más general, enriqueciéndolo con nuevos escenarios.

Supóngase que para un primer escenario, la calibración correspondiente que se efectúa da como resultado un valor $\lambda = \lambda_0$. Ese cálculo garantiza que la tasa de discrepancias Δ cumple los requisitos

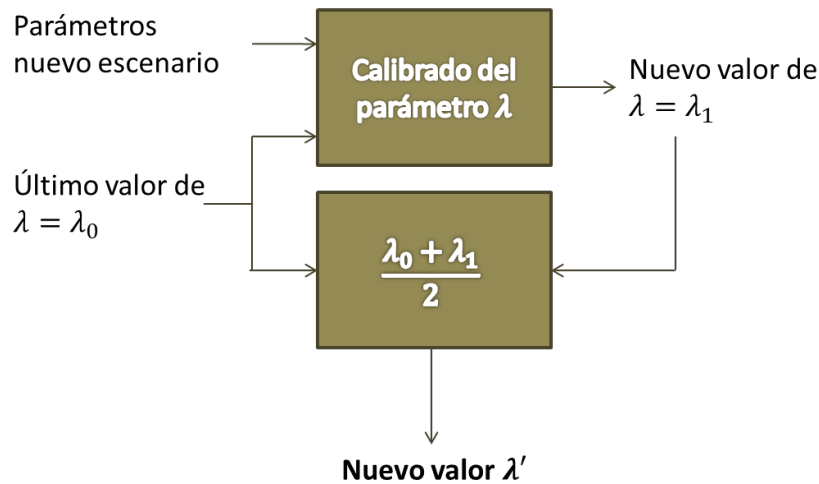


Figura 8.9: Marco de implementación

establecidos. Ahora se desea introducir un segundo escenario en el modelo. Introduciendo los datos del nuevo escenario (nuevo tráfico, nuevos recursos) y aprovechando el parámetro $\lambda = \lambda_0$ anterior, se obtiene un nuevo valor de la tasa de discrepancias Δ' . Se ha producido una variación en la señal de salida (tasa de discrepancias) y es en ese momento donde se pone en marcha el sistema de calibración para determinar un nuevo valor de λ que satisfaga en la medida de lo posible la calidad para ambos escenarios. Si en el nuevo recálculo se ha determinado que, con los recursos disponibles y el objetivo propuesto, $\lambda = \lambda_1$, entonces se realizará el recálculo de la accesibilidad con un nuevo parámetro λ' que recoja de alguna manera los escenarios anteriores con los que se ha calibrado el sistema, de acuerdo con la ecuación 8.4.

$$\lambda' = \frac{\lambda_0 + \lambda_1}{2} \quad (8.4)$$

Este procedimiento se repetirá tantas veces como escenarios de entrenamiento se consideren, esperando encontrar un valor estable.

Para la recalibración del sistema se recurre al sistema realimentado descrito en el apartado anterior. El método de prueba iterativa para valores de λ crecientes proporciona la solución correctamente; no obstante, y dado que los cálculos de accesibilidad son computacionalmente costosos, se diseña el siguiente procedimiento (ver figura 8.10) para favorecer una búsqueda rápida del nuevo valor de λ .

1. Se parte del último valor de λ calibrado según las experiencias anteriores, $\lambda = \lambda_0$.
2. Se introducen los datos del nuevo escenario (tráfico reportado, recursos disponibles...) en el modelo de que se dispone y se comprueba la validez/calidad del mismo.
 - a) Si el sistema no cumple con los criterios de calidad, se busca un nuevo valor de λ . Como se ha concluido en el capítulo previo, mayores valores de λ dan como resultado una menor tasa de discrepancias, se decide incrementar su valor en 0.5. Así, $\lambda_0^1 = \lambda_0 + 0,5$. Con este parámetro se continúa en el paso (3).
 - b) Si el sistema cumple los criterios de calidad, también se buscará un nuevo valor de λ . Es sabido que valores más pequeños dan como resultado un menor error de inventariado,

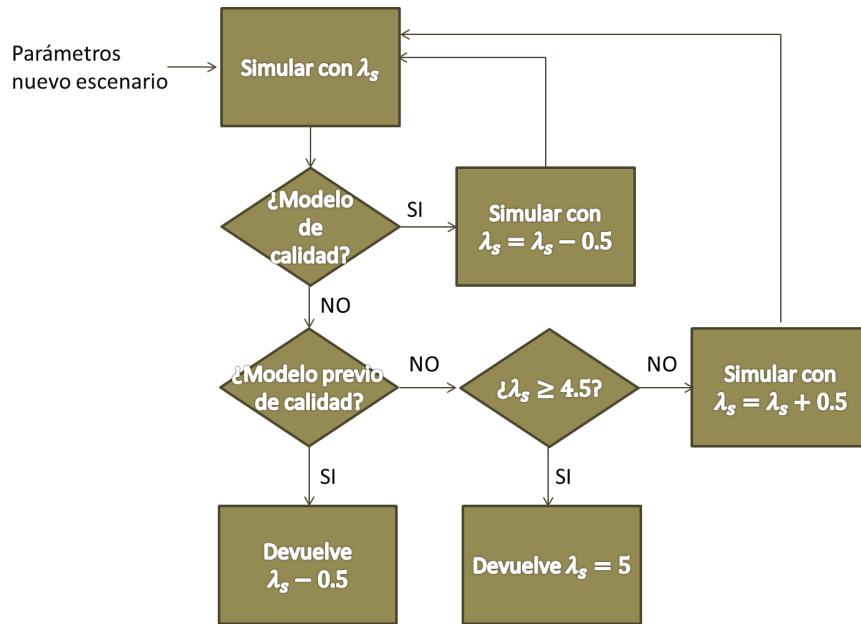


Figura 8.10: Procedimiento para la verificación de calidad vs. λ

por lo que se decide probar qué sucedería con un valor más pequeño, $\lambda_0^1 = \lambda_0 - 0,5$. Con este parámetro se continúa en el paso (3).

3. Se modifica el modelo con el nuevo parámetro λ_0^1 y se comprueba la calidad del mismo con los datos del nuevo escenario que se está calibrando.

a) Si el sistema no cumple los criterios de calidad, y antes tampoco los cumplía, se continúa incrementando el valor de λ con un nuevo salto de 0,5, por lo que $\lambda_0^2 = \lambda_0^1 + 0,5$. Como se ha establecido una cota superior para λ para controlar el error de inventariado:

1) Si se alcanza ese valor, se detiene la búsqueda y se devuelve $\lambda_1 = 5$ con el «warning» de que el sistema no cumple los requisitos de calidad.

2) Si aún no se alcanza ese valor, se vuelve al paso (2a) con nuevos parámetros λ_0^n .

b) Si el sistema no cumple los criterios de calidad, pero antes sí que los cumplía, se detiene la búsqueda y se resuelve que el mejor valor de λ era el anterior, $\lambda_1 = \lambda_0^{n-1}$.

c) Si el sistema cumple los criterios de calidad, se continúa en el paso (2b).

4. El nuevo valor λ_1 se combina con el parámetro inicial λ_0 para tener en cuenta las experiencias anteriores, de acuerdo con la expresión 8.4. El nuevo valor λ' se convierte en el nuevo λ_0 para futuras calibraciones con nuevos escenarios.

De acuerdo con los resultados que se han obtenido hasta el momento, a la hora de diseñar el subsistema de interpolación del tráfico existe un compromiso a la hora de evaluar λ . En efecto, valores más pequeños de λ provocan una mayor fiabilidad de reconstrucción, pero al mismo tiempo valores más elevados proporcionan una menor tasa de discrepancias Δ . Por ello, y por coherencia con el contexto en el que se desarrolla la Tesis, es imprescindible fijar una cota superior para λ dentro del sentido común. Se elige como cota $\lambda = 5$ por presentar una situación aceptable desde ambos puntos de vista. El lector puede consultar la implementación del código fuente en Matlab en el apartado A.4.5.

8.4. Evaluación de los resultados

Tanto para el proceso de calibrado como de validación del modelo que se ha desarrollado en esta Tesis, se han utilizado nodos B de un operador móvil real. Se utilizan por tanto estadísticas de tráfico cursado reales, lo cual supone una importante contribución con respecto a otras investigaciones encontradas, basadas únicamente en simulaciones.

Gracias a los KPIs reportados por estos nodos, se ha podido validar este modelo en términos de su accesibilidad global. Así, su calidad se mide de acuerdo con su capacidad de detectar la congestión y más concretamente de identificar cuáles son los recursos limitantes. Para ello, se parte del análisis de accesibilidad de un nodo en 2 situaciones:

- ⇒ En un primer momento, se analiza la capacidad del modelo para detectar la degradación, teniendo un recurso subdimensionado.
- ⇒ Posteriormente, se amplía este recurso y se verifica si el modelo detecta una disminución del bloqueo en términos generales y más en concreto, del factor antes limitante.

Así, el proceso de estudio de estos nodos (antes y después) se prolonga durante varios días consecutivos con demandas de tráfico similares.

Además, como se verá posteriormente, este modelo permite ayudar al operador de red móvil con el que se ha trabajado a discriminar entre bloqueo por congestión y fallos de tipo hardware. A tal efecto, se ha validado también el modelo utilizando un nodo real del que sabía que tenía fallos hardware.

De acuerdo con el modelo implementado, para cada nodo es necesario conocer:

- ⇒ Capacidad total disponible de cada recurso.
- ⇒ Servicios involucrados (desde voz a HSDPA).
- ⇒ Intensidad de tráfico e intentos de conexión (obtenidos a partir de las estadísticas reales de los nodos).

8.4.1. Nodo «A»: nodo con capacidad IUB limitada (escenarios 1 y 2)

Se trata de un nodo con configuración DualStack del que se analizarán dos escenarios diferentes, modificando alguno de sus componentes de un escenario a otro.

8.4.1.1. Escenario 1

Está equipado con los siguientes recursos:

- ⇒ CE UL: 768 CE.
- ⇒ CE DL: 768 CE.
- ⇒ SC: 2 portadoras.
- ⇒ Iub UL: 1 ATM VCC.
- ⇒ Iub DL: IP 23 Mbps.

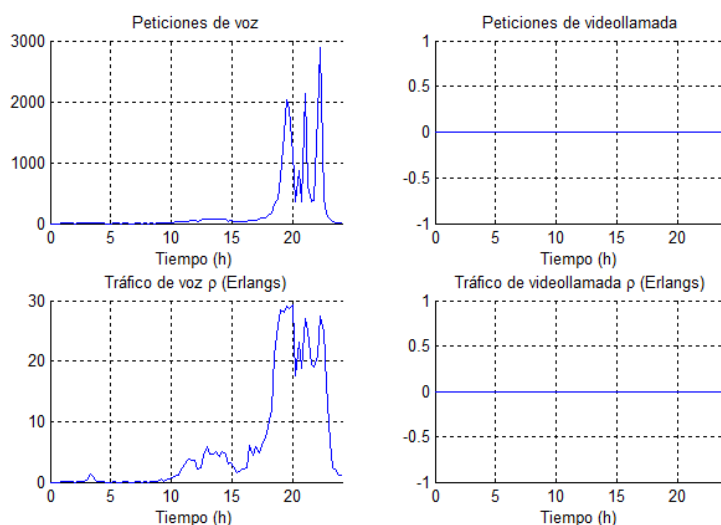


Figura 8.11: Tráfico CS durante un día de observación (nodo A, antes)

Se dispone de las estadísticas de un día completo que, como ya se ha comentado, tienen una granularidad de 15 minutos, la cual se modelará para obtener una granularidad más efectiva, de 1 minuto. Los datos de peticiones y de tráfico en el nodo, para el día estudiado, se pueden ver en las figuras 8.11 y 8.12.

En la figura 8.13 se presentan la accesibilidades teórica y real para tráfico rígido.

En la parte superior de la figura puede verse la evolución temporal de la accesibilidad real junto con la calculada teórica. Al observar ambas, se corrobora que existe una degradación elevada, situación que se reproduce en la curva teórica resultante del modelo. Numéricamente los valores de accesibilidad no coinciden exactamente, debido fundamentalmente a que en el modelo no se han tenido en cuenta algunos factores adicionales que pueden influir en la accesibilidad, por ejemplo las condiciones radio, pérdidas de propagación, interferencias, limitaciones de potencia, etc. A pesar de la diferencia cuantitativa, el modelo reproduce la forma y cumple su propósito, puesto que permite detectar cuándo se produce la degradación.

Más aún, el modelo permite determinar cuál es el recurso por el cual el nodo sufre más bloqueo. Así, en la parte central e inferior de la figura 8.13 se representan las accesibilidades parciales referidas a cada uno de los recursos estudiados (DL y UL, respectivamente). Se observa que el recurso causante en mayor medida de los bloqueos es el correspondiente al IUB.

Por otra parte, en la figura 8.14 se representan los mismos resultados pero ahora para el tráfico PS. Nuevamente el modelo prevé correctamente la degradación, siendo el recurso más degradado los códigos de canalización.

Para estudiar desde un punto de vista más exacto la capacidad que tiene el modelo de predecir la degradación se recurre a una digitalización de la señal, transformando las curvas anteriores de accesibilidad en series binarias de 0s y 1s, donde 0 significa que hay degradación y 1 que no la hay.

En la figura 8.15 se presentan los resultados de esta digitalización. Para facilitar la visualización por parte del lector, se digitalizan con el alfabeto (0,1) la accesibilidad real y (0,-1) la teórica, para que de este modo puedan compararse fácilmente las predicciones. Como puede observarse, salvo errores puntuales, el modelo detecta fielmente la degradación en el sistema.

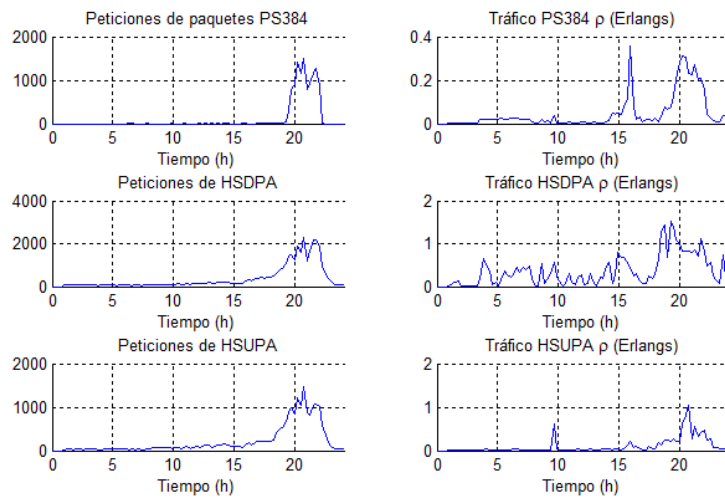


Figura 8.12: Tráfico PS durante un día de observación (nodo A, antes)

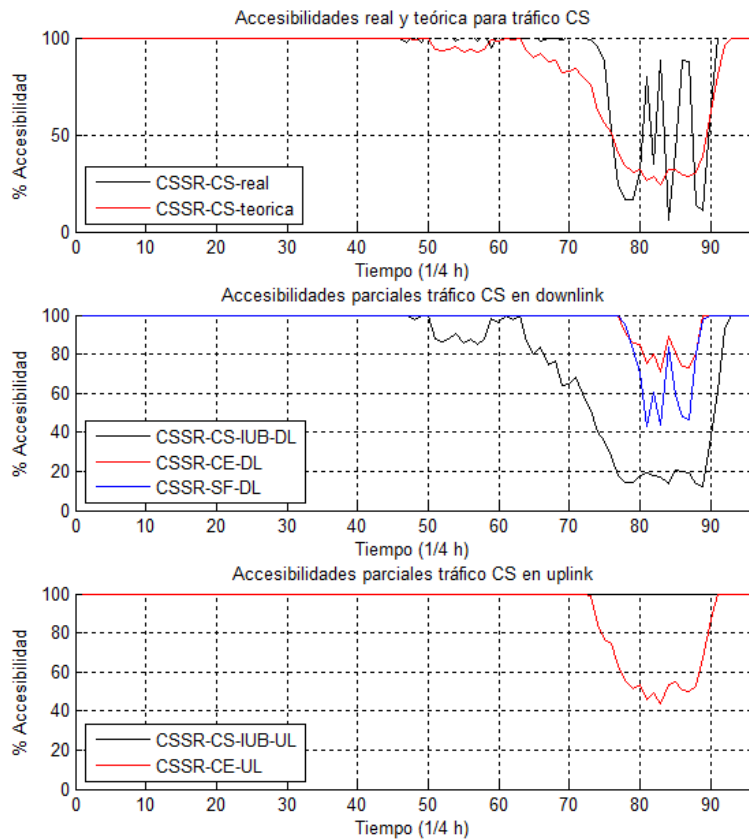


Figura 8.13: Accesibilidades real y teórica para tráfico CS en el nodo A.

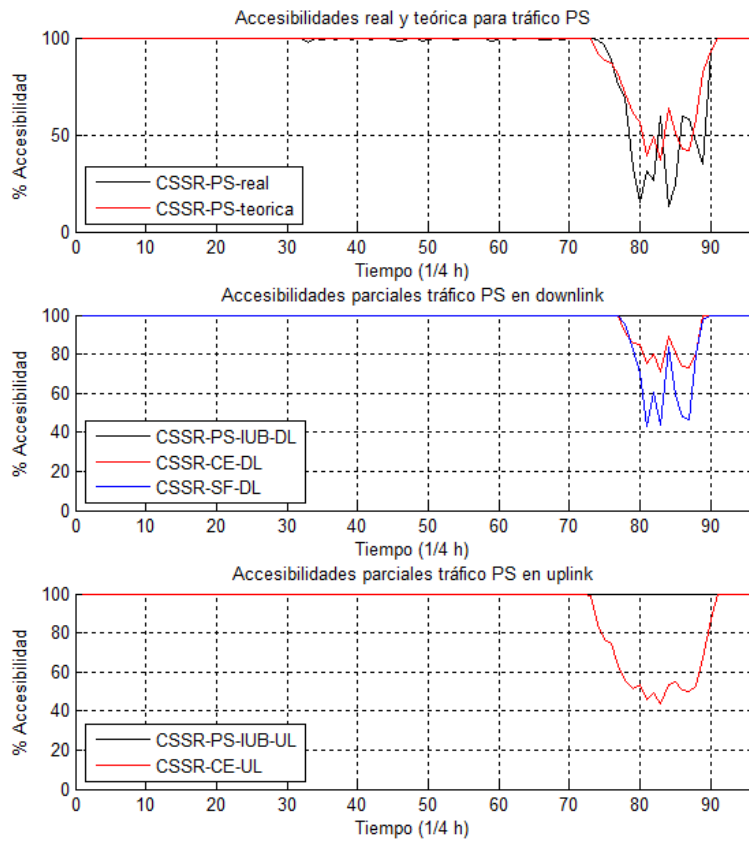


Figura 8.14: Accesibilidades real y teórica para tráfico PS en el nodo A.

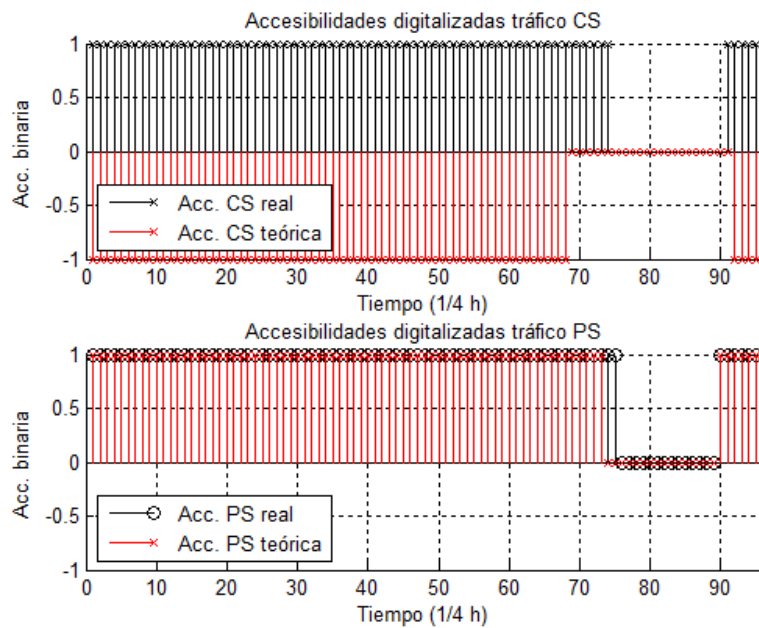


Figura 8.15: Predicción de la degradación en modelos teórico y real (nodo A, antes)

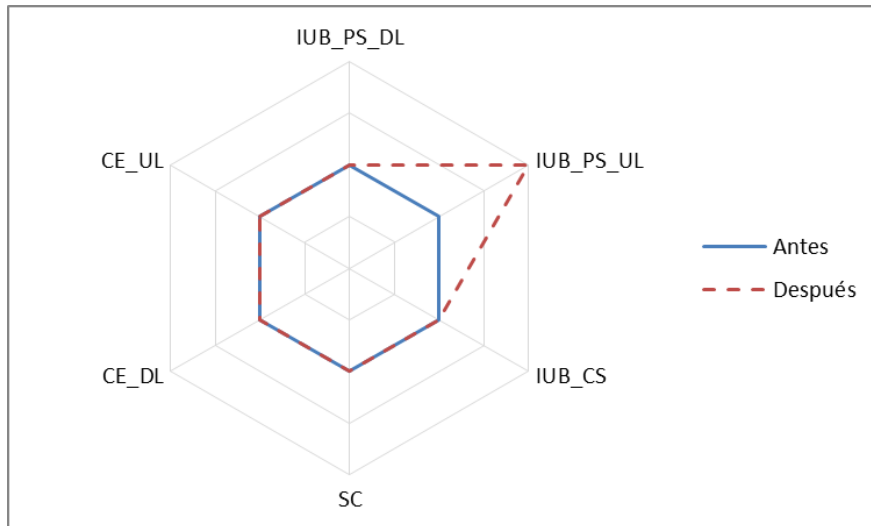


Figura 8.16: Ampliación de recursos en el nodo «A»

8.4.1.2. Escenario 2

Debido a la alta congestión y baja accesibilidad que presentaba el escenario 1 se decide realizar una reasignación de recursos. Se aumenta la capacidad disponible en el Iub para servicios rígidos pasando a tener 2 VCC en vez de uno (figura 8.16).

Esto debería de mejorar la accesibilidad en el Iub, lo cual se comprobará a continuación. Finalmente, los recursos del nodo son:

⇒ CE UL: 768 CE.

⇒ CE DL: 768 CE.

⇒ SC: 2 portadoras.

⇒ Iub UL: 2 ATM VCC.

⇒ Iub DL: IP 23 Mbps.

En las figuras 8.17 y 8.18 se muestran los datos de entrada para un día de observación.

En las figuras 8.19 y 8.20 se representan, respectivamente, los resultados de accesibilidad con el recurso IUB ampliado. En ambas se observa que los valores de accesibilidad han mejorado respecto al escenario anterior. Además, a la vista de las mismas se puede concluir que se trata de un nodo con problemas de congestión, ya que la accesibilidad real es más alta que la teórica.

Por último, en la figura 8.21 se analiza el rendimiento del modelo ante esta nueva situación. Como puede observar el lector, salvo errores puntuales, en la mayoría de los casos el modelo es capaz de predecir la degradación real.

8.4.2. Nodo «B» (escenarios 3 y 4)

Se trata también de un nodo con configuración DualStack en el que se estudiarán dos escenarios.

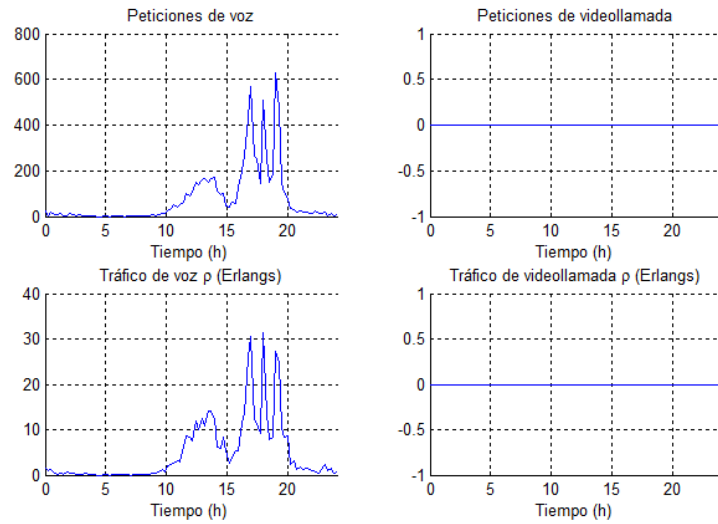


Figura 8.17: Tráfico CS durante un día de observación (nodo A, después)

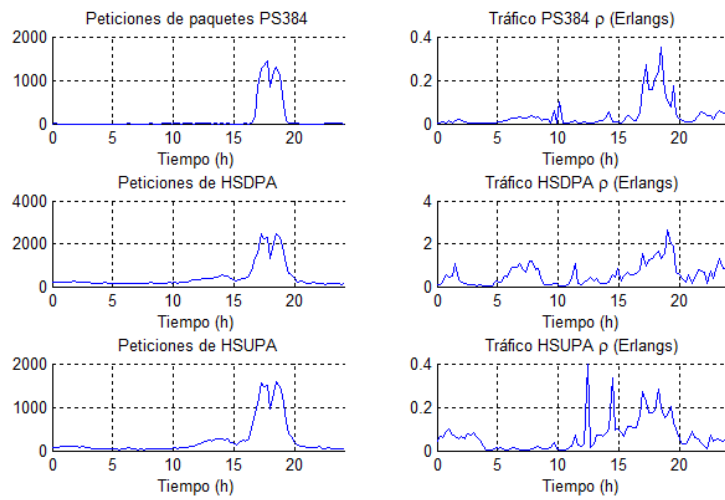


Figura 8.18: Tráfico PS durante un día de observación (nodo A, después)

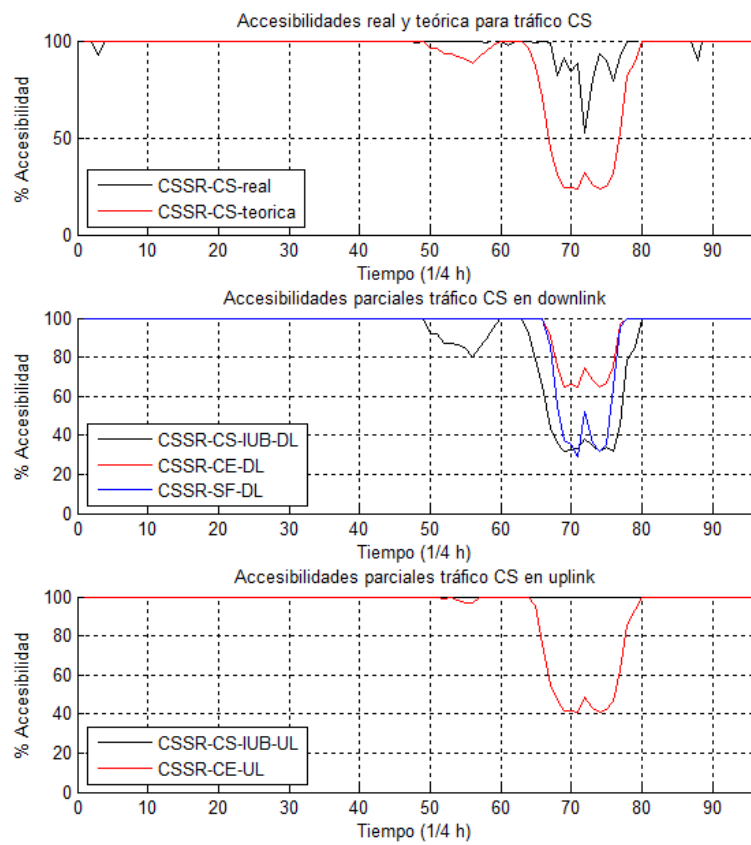


Figura 8.19: Accesibilidades real y teórica para el tráfico CS en el nodo A después de incrementar el recurso IUB.

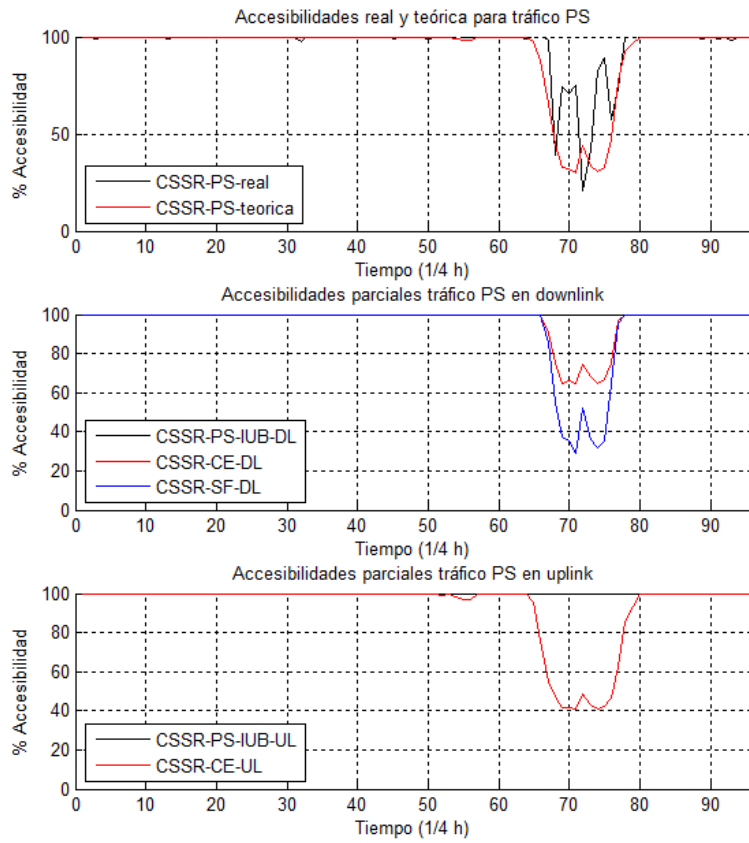


Figura 8.20: Accesibilidades real y teórica para el tráfico PS en el nodo A después de incrementar el recurso IUB

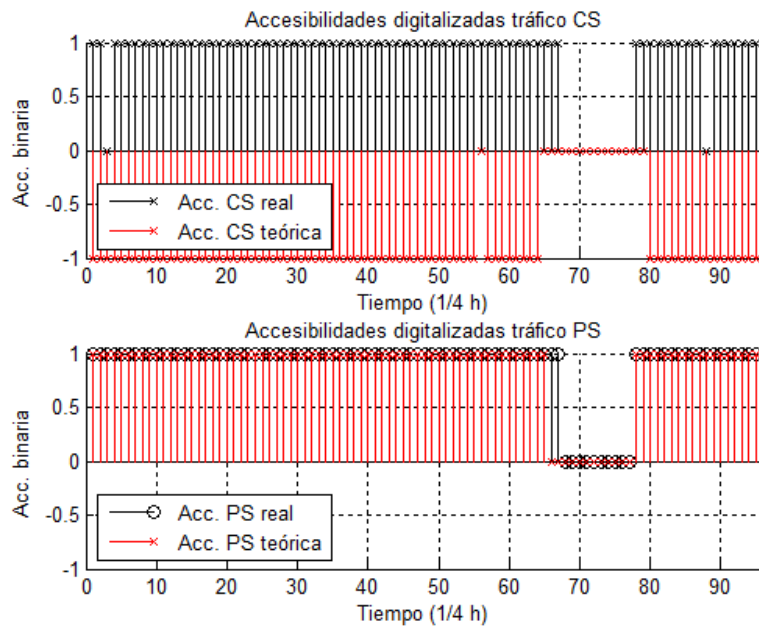


Figura 8.21: Predicción de la degradación en modelos teórico y real (nodo A, después)

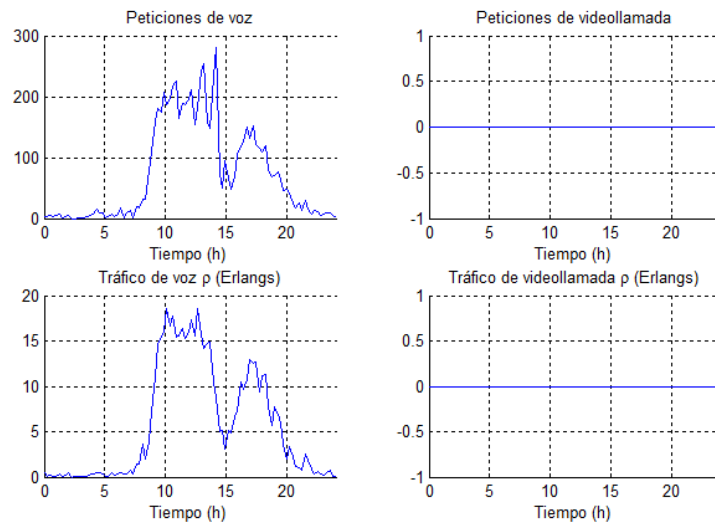


Figura 8.22: Tráfico CS durante un día de observación (nodo B, antes)

8.4.2.1. Escenario 3

Las características de este escenario, son iguales que las del nodo anterior, es decir, se trata de un nodo dotado de los siguientes recursos:

- ⇒ CE UL: 768 CE.
- ⇒ CE DL: 768 CE.
- ⇒ SC: 2 portadoras.
- ⇒ Iub UL: 1 ATM VCC.
- ⇒ Iub DL: IP 23 Mbps.

Se trata también de un nodo en servicio, correspondiente a un operador real, cuyas estadísticas se analizan para un día completo. En las figuras 8.22 y 8.23 se muestran las peticiones y los tráficos de entrada para el modelo:

A partir de dichas gráficas, puede observarse que la congestión se concentra en intervalos de tiempo muy localizados. En las figuras 8.24 y 8.31 se muestran las accesibilidades correspondientes al tráfico rígido y elástico, respectivamente. En la parte superior, comparativamente la accesibilidad real y la teórica, en la central, el desglose de accesibilidad teórica en cada uno de los posibles factores limitantes para el enlace DL y en la parte inferior las accesibilidades parciales para el enlace UL.

En este caso, el factor más limitante en el enlace descendente es la capacidad del Iub de circuitos, seguido por los Channel Elements. Debido a esto se optará por aumentar la capacidad del Iub pasando a tener 2 VCC en el escenario 4. En el enlace ascendente, en cambio, el factor más limitante son los Channel Elements, pero se consigue obtener una accesibilidad por encima del 80%. En este enlace no afecta la capacidad del Iub de paquetes y la de circuitos tiene una accesibilidad muy alta.

En cuanto a la accesibilidad total en este escenario, se distinguen dos tramos de congestión, tanto para circuitos como para paquetes. Aun así, la accesibilidad de paquetes es un poco mejor que

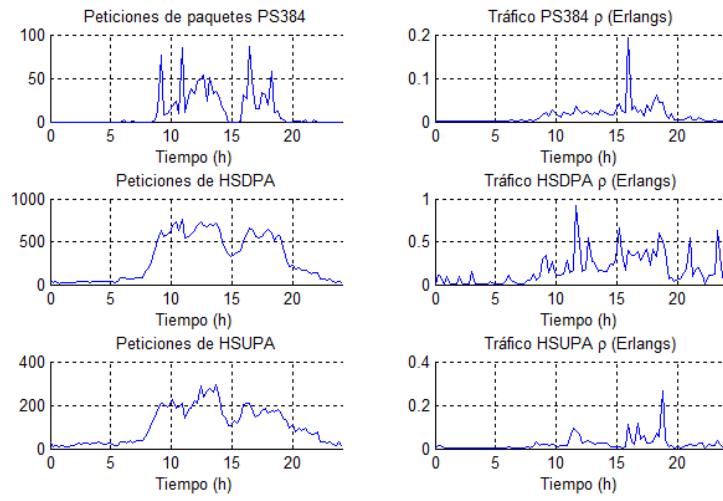


Figura 8.23: Tráfico PS durante un día de observación (nodo B, después)

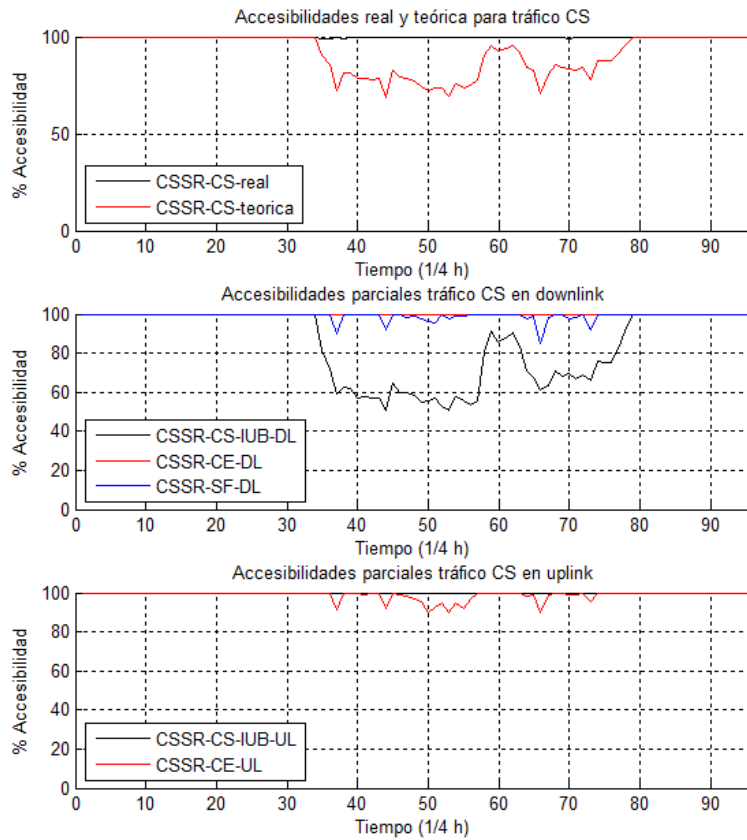


Figura 8.24: Accesibilidades real y teórica para el tráfico CS en el nodo B (antes)

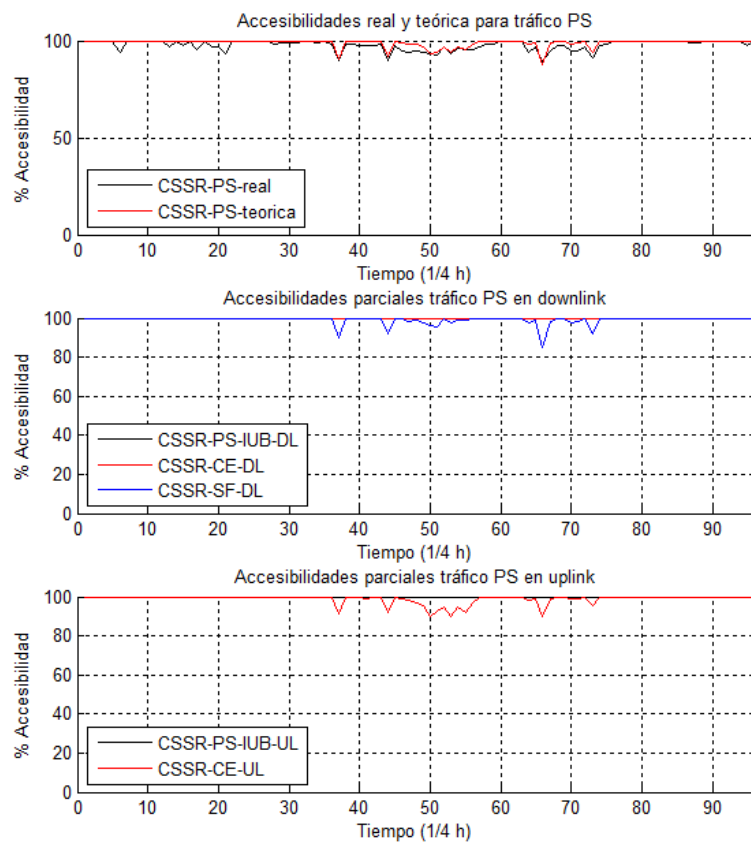


Figura 8.25: Accesibilidades real y teórica para el tráfico PS en el nodo B (antes)

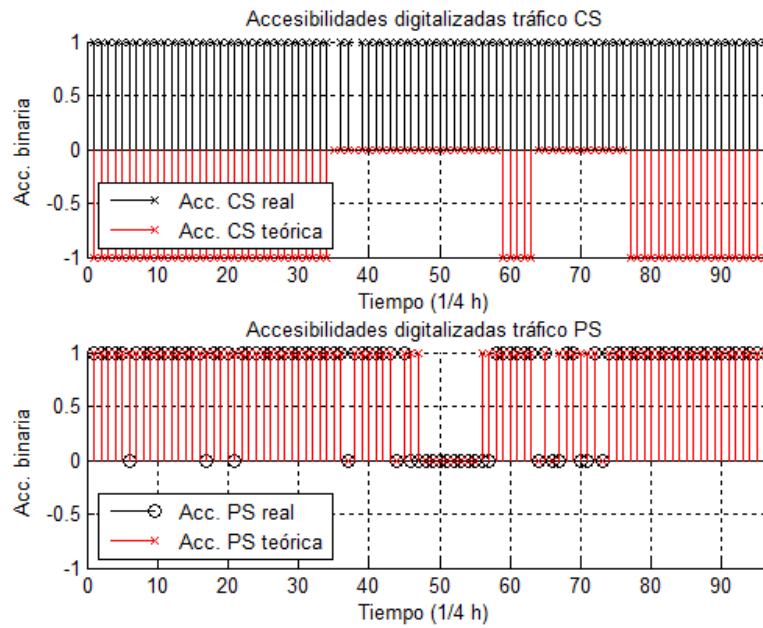


Figura 8.26: Predicción de la degradación en modelos teórico y real (nodo B, antes)

la accesibilidad de circuitos. Como la accesibilidad teórica es claramente menor que la accesibilidad real se trata de un problema de congestión.

Para comprobar una vez más la fiabilidad del modelo a la hora de detectar o no la degradación, se comparan las curvas de accesibilidades digitalizadas, del mismo nodo que con el nodo anterior. En la figura 8.26 se aprecian los resultados.

8.4.2.2. Escenario 4

Al igual que sucedida en el nodo A, en este nodo B también se aumenta la capacidad del Iub pasando a tener 2 VCC (figura 8.27). Así, los recursos de que dispone el nodo son:

- ⇒ CE UL: 768 CE.
- ⇒ CE DL: 768 CE.
- ⇒ SC: 2 portadoras.
- ⇒ Iub UL: 2 ATM VCC.
- ⇒ Iub DL: IP 23 Mbps.

A continuación, en las figuras 8.28 y 8.29 se muestran las gráficas de los datos de entrada para este escenario.

Introduciendo los datos anteriores, así como el equipamiento del nodo, como parámetros de entrada al modelo, se obtienen las accesibilidades teóricas tanto para el tráfico rígido (figura 8.30) como para el elástico (figura 8.31).

Comprando las accesibilidades de los escenarios 3 y 4, se comprueba que al aumentar la capacidad del Iub pasando de 1 VCC a 2 VCC se ha mejorado la accesibilidad por Iub, pasando a ser ahora los Channel Elements el factor limitante.

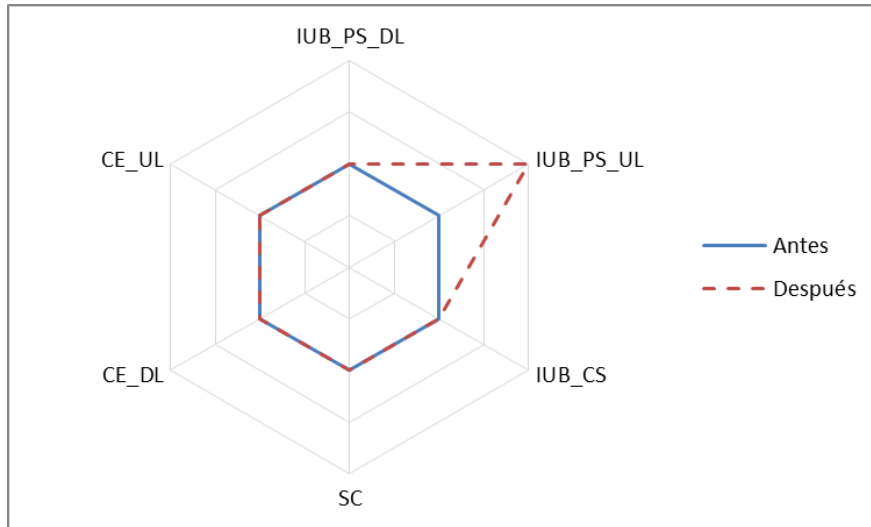


Figura 8.27: Ampliación de recursos en el nodo «B»

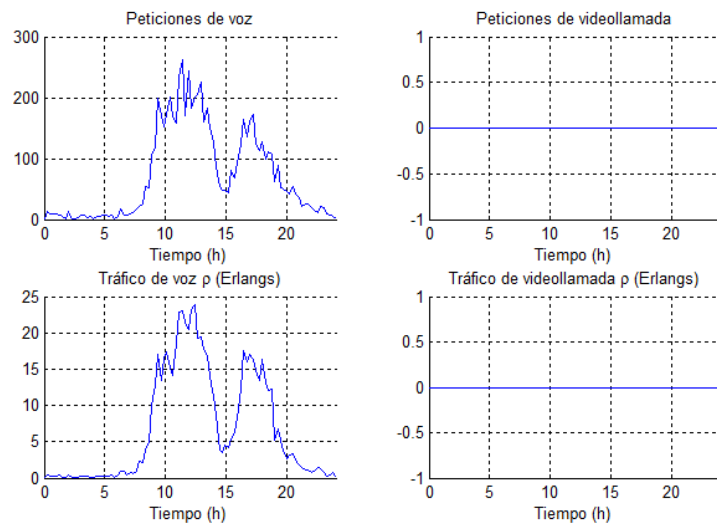


Figura 8.28: Tráfico CS durante un día de observación (nodo B, después)

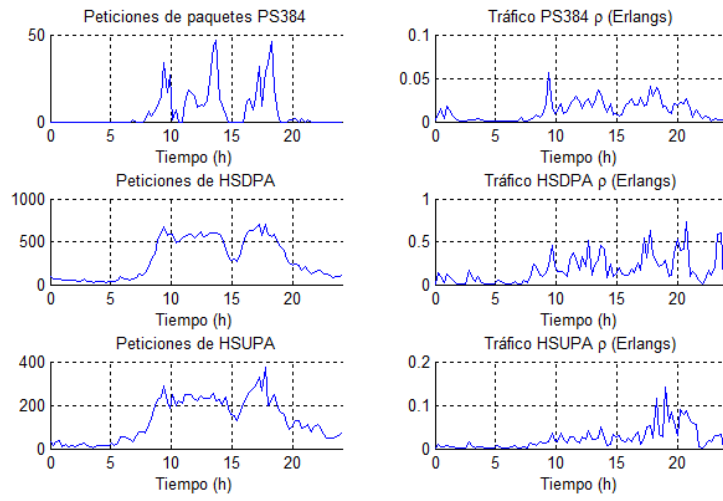


Figura 8.29: Tráfico PS durante un día de observación (nodo B, después)

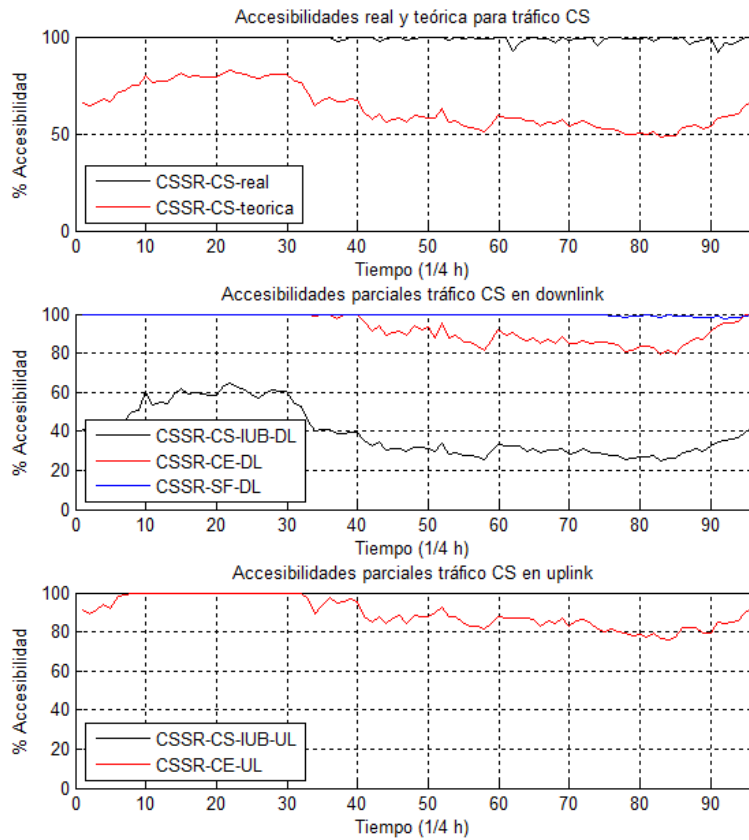


Figura 8.30: Accesibilidades real y teórica para el tráfico CS en el nodo B (después)

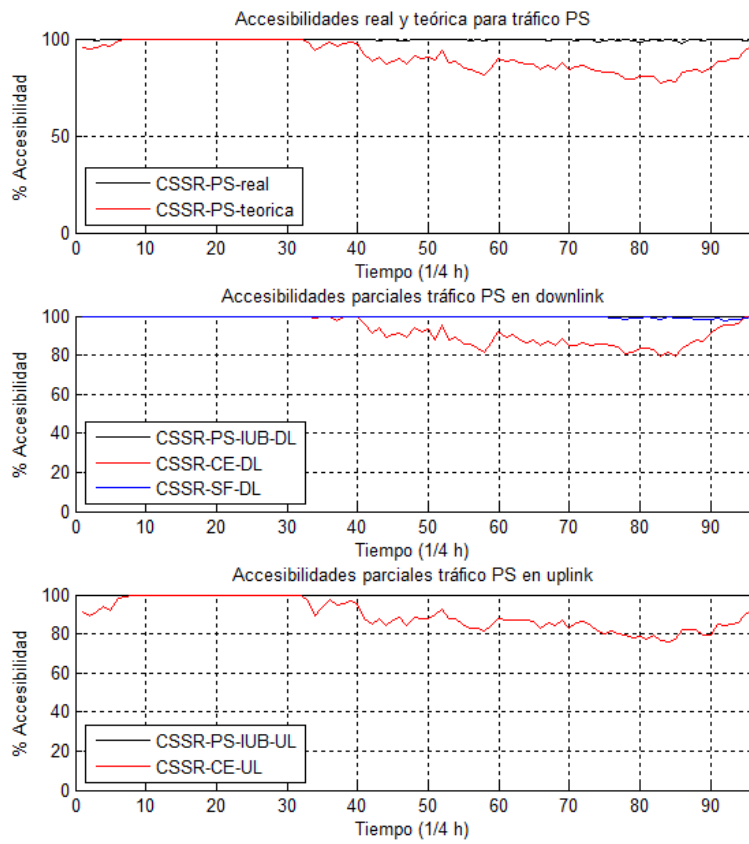


Figura 8.31: Accesibilidades real y teórica para el tráfico PS en el nodo B (después)

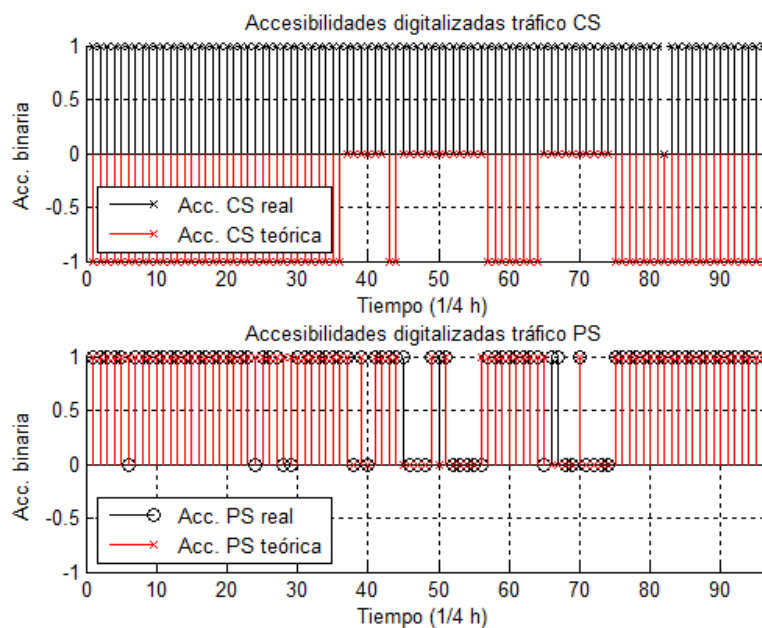


Figura 8.32: Predicción de la degradación en modelos teórico y real (nodo B, después)

A su vez, al mejorar la accesibilidad por Iub se mejora también la accesibilidad total, como se puede ver comparando las gráficas anteriores con las del escenario 3. Por ejemplo, se ha pasado de tener una accesibilidad de circuitos entre el 50-70 % a una entre el 70-80 %. El siguiente paso para mejorar la accesibilidad de este escenario sería aumentar la capacidad de Channel Elements, ya que es el siguiente factor limitante.

Por último, para verificar una vez más la fiabilidad del modelo, se procede a digitalizar ambas señales a fin de comprobar coincidencias (0,0) y (1,-1). Los resultados se ilustran en la figura 8.32.

8.4.3. Nodo «C», ampliación del recurso CE (escenarios 5 y 6)

8.4.3.1. Escenario 5

Los recursos con los que queda equipado el nodo son:

⇒ CE UL: 768 CE.

⇒ CE DL: 768 CE.

⇒ SC: 1 portadora.

⇒ Iub UL: 1 ATM VCC.

⇒ Iub DL: IP 30 Mbps.

A continuación, en las figuras 8.33 y 8.34 se muestran las gráficas de los datos de entrada para este escenario.

A la vista de las figuras 8.35 y 8.36, todos los factores limitan casi por igual, a excepción de la capacidad Iub para el tráfico de paquetes, que no es limitante. Debido a la baja accesibilidad por Channel Elements en el enlace ascendente se intentará mejorar este factor añadiendo otra tarjeta en el enlace ascendente (CE_UL). Al comparar la accesibilidad teórica y la real, tanto de circuitos

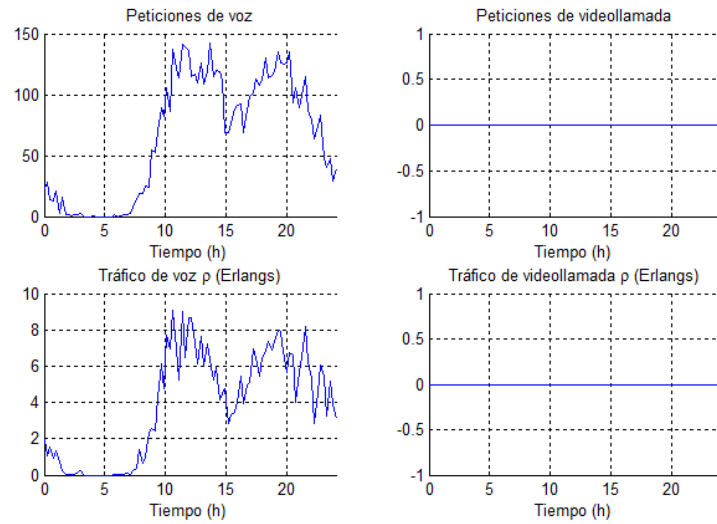


Figura 8.33: Tráfico CS durante un día de observación (nodo C, antes)

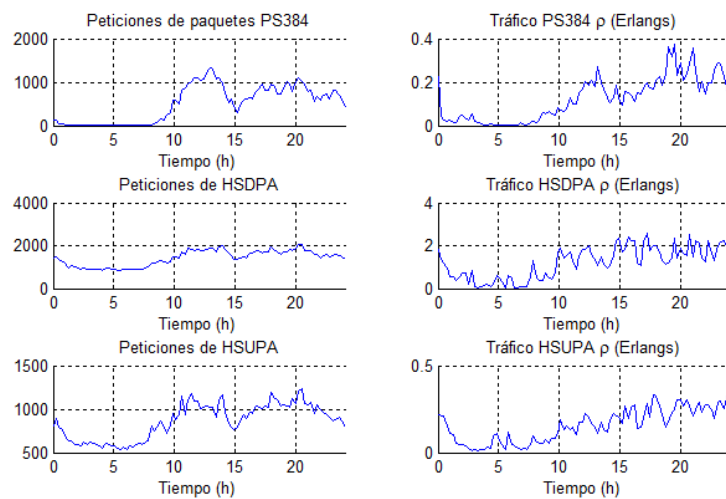


Figura 8.34: Tráfico PS durante un día de observación (nodo C, después)

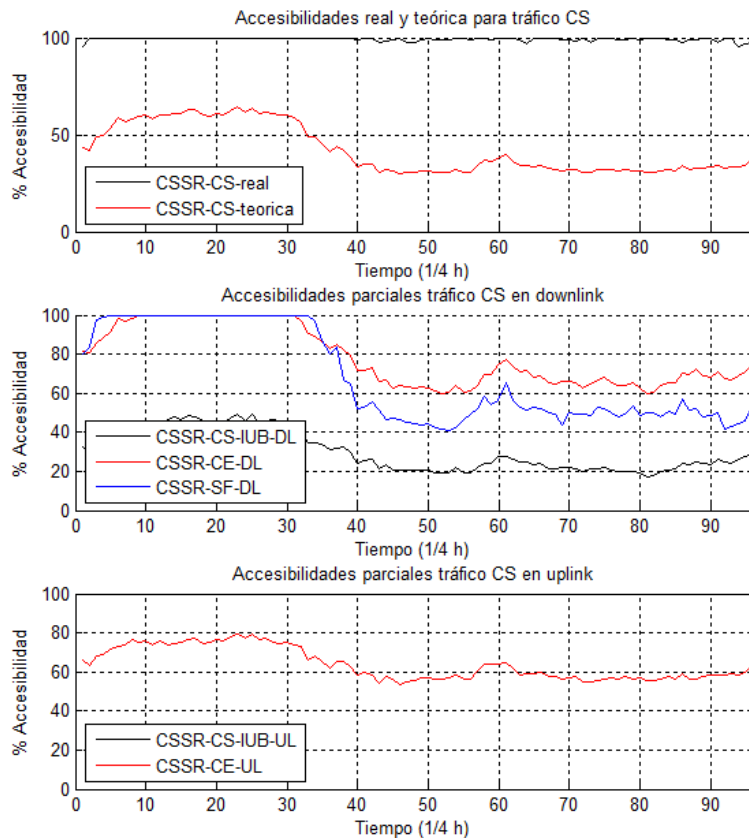


Figura 8.35: Accesibilidades real y teórica para el tráfico CS en el nodo C (antes)

como de paquetes se aprecia que la accesibilidad real es mayor lo que implica un problema de congestión ya que existen más peticiones bloqueadas teóricas.

Por último, en la figura 8.37 se presentan las accesibilidades binarias para tráfico CS y PS, lo que permite comprobar el comportamiento predictivo del modelo, del mismo modo que en escenarios anteriores. Mientras que la predicción para tráfico elástico es prácticamente perfecta, para el tráfico CS se aprecian un mayor número de discrepancias. El carácter de estas muestras reside en que la curva de accesibilidad teórica presenta un comportamiento significativamente más «pesimista» que la real, que en el proceso de digitalización refleja que las predicciones pueden depender en algunos casos de la elección del umbral de discretizado.

8.4.3.2. Escenario 6

Debido a la congestión del nodo, y partiendo de lo estudiado en los nodos A y B, en los que se ha comprobado que uno de los factores más limitantes son los Channel Elements, en este caso se opta por aumentar su número (figura 8.38). Particularmente se aumenta el número de tarjetas en el enlace ascendente, pasando de tener 2 tarjetas de CE_UL a tener 3. El nodo queda configurado de la siguiente forma:

⇒ CE UL: 1152 CE.

⇒ CE DL: 768 CE.

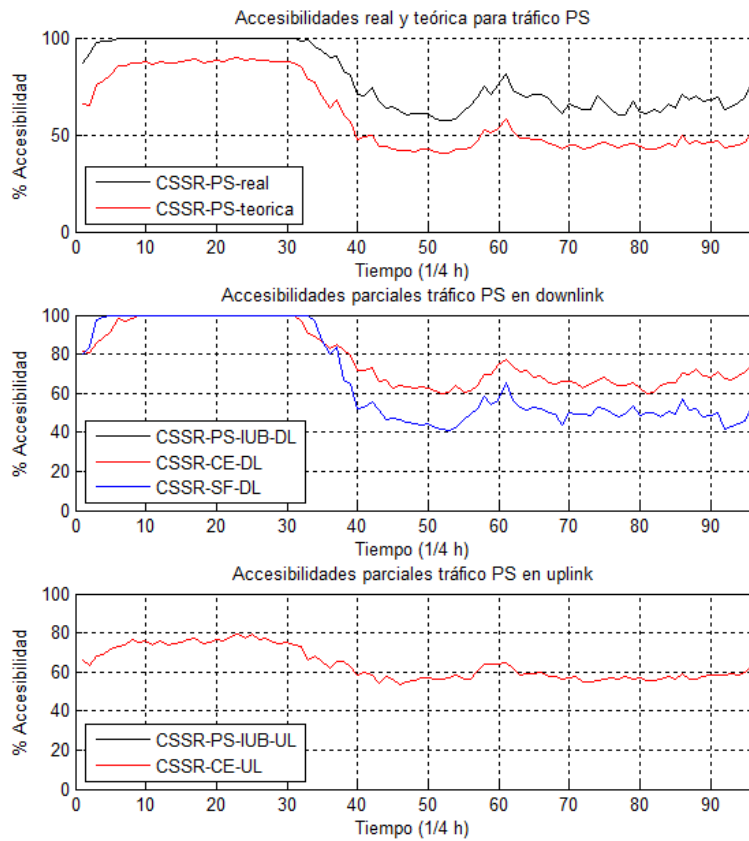


Figura 8.36: Accesibilidades real y teórica para el tráfico PS en el nodo C (antes)

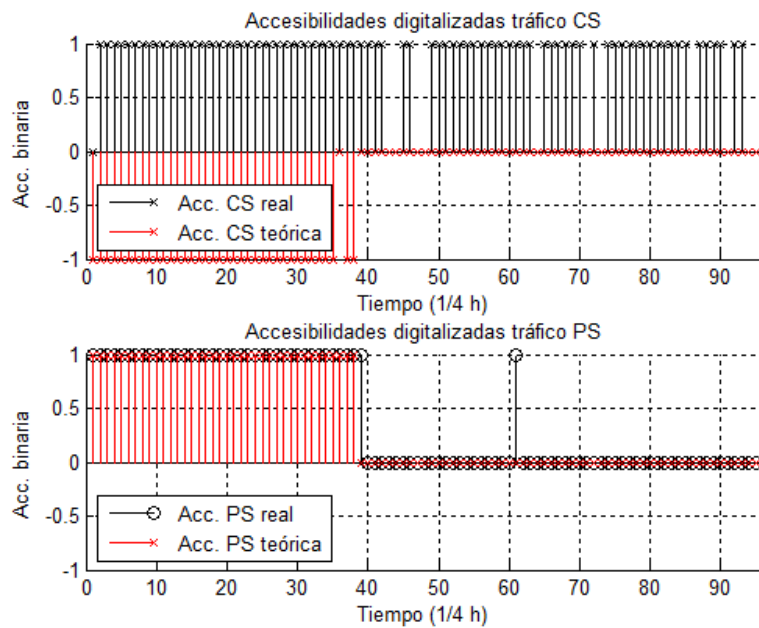


Figura 8.37: Predicción de la degradación en modelos teórico y real (nodo C, antes)

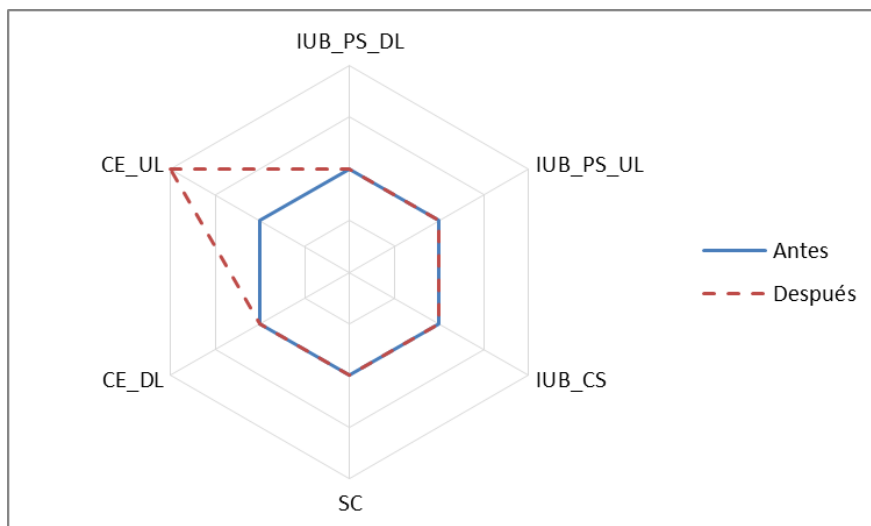


Figura 8.38: Ampliación de recursos en el nodo «C»

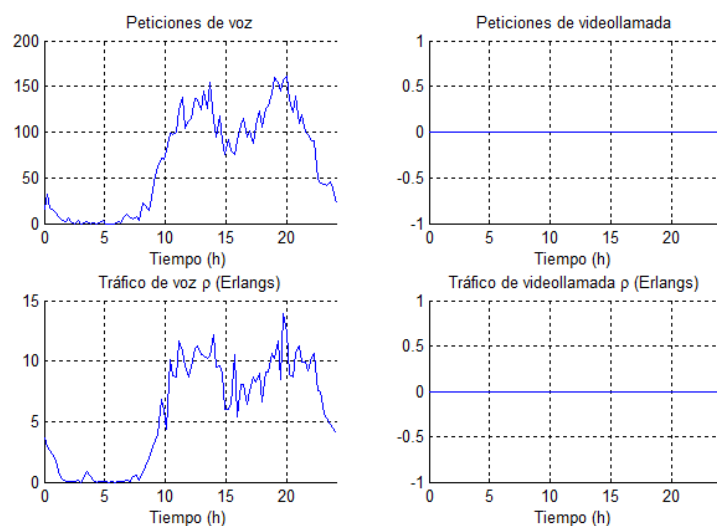


Figura 8.39: Tráfico CS durante un día de observación (nodo C, después)

⇒ SC: 1 portadora.

⇒ Iub UL: 1 ATM VCC.

⇒ Iub DL: IP 30 Mbps.

Esta equipación, junto con el tráfico registrado por el nodo y presentado en las figuras 8.39 y 8.40, constituyen los parámetros de entrada de este escenario en el modelo.

En las figuras 8.41 y 8.42 se representan las accesibilidades teóricas para tráficos CS y PS. Se puede ver una gran mejora en la accesibilidad por Channel Elements comparando el escenario anterior. Esto es debido, no solo al aumento del número de Channel Elements, si no a que ha disminuido el tráfico: hay un considerable menor número de peticiones HSUPA, lo que ayuda a la gran mejora que experimentamos en este nuevo escenario. Todo esto tiene como consecuencia que también mejore la accesibilidad total de circuitos y paquetes, como se puede apreciar en las

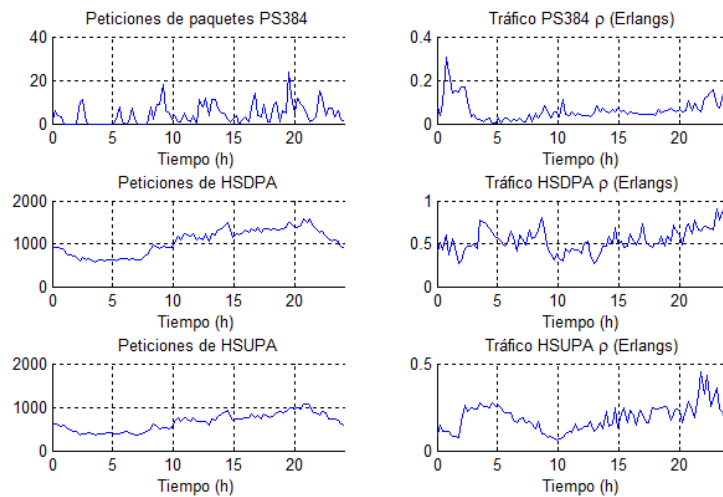


Figura 8.40: Tráfico PS durante un día de observación (nodo C, después)

figuras.

Por último, como en el resto de escenarios, en la figura 8.43 se representa la digitalización de las accesibilidades teóricas, a fin de evaluar la capacidad de predicción del modelo. Se observa nuevamente una clara dependencia del umbral de discriminación que se toma en el proceso de digitalización.

8.4.4. Nodo «D», configuración Full-IP (escenarios 7 y 8)

En este caso, y a diferencia de los anteriores, se trata de un nodo con configuración Full-IP, por lo que toda la transmisión se realiza utilizando la capacidad de la conexión ADSL, creando un VCC virtual. Al igual que en el resto de nodos se analizarán 2 escenarios.

8.4.4.1. Escenario 7

Las características del primer escenario para este nodo son:

- ⇒ CE UL: 768 CE.
- ⇒ CE DL: 768 CE.
- ⇒ SC: 1 portadora.
- ⇒ Iub UL: 30 Mbps.
- ⇒ Iub DL: 10 Mbps.

En las figuras 8.50 y 8.51 se representan los tráficos CS y PS que ha cursado este nodo, respectivamente.

Los datos anteriores son introducidos en el modelo de acuerdo a las especificaciones particulares presentadas en el capítulo anterior. Las accesibilidades teóricas obtenidas se representan en las figuras 8.46 y 8.47.

Se aprecia en las gráficas anteriores que el elemento limitante son los Channel Elements, como en varios de los casos anteriores. El segundo factor más limitante son los Spreading Codes, mientras

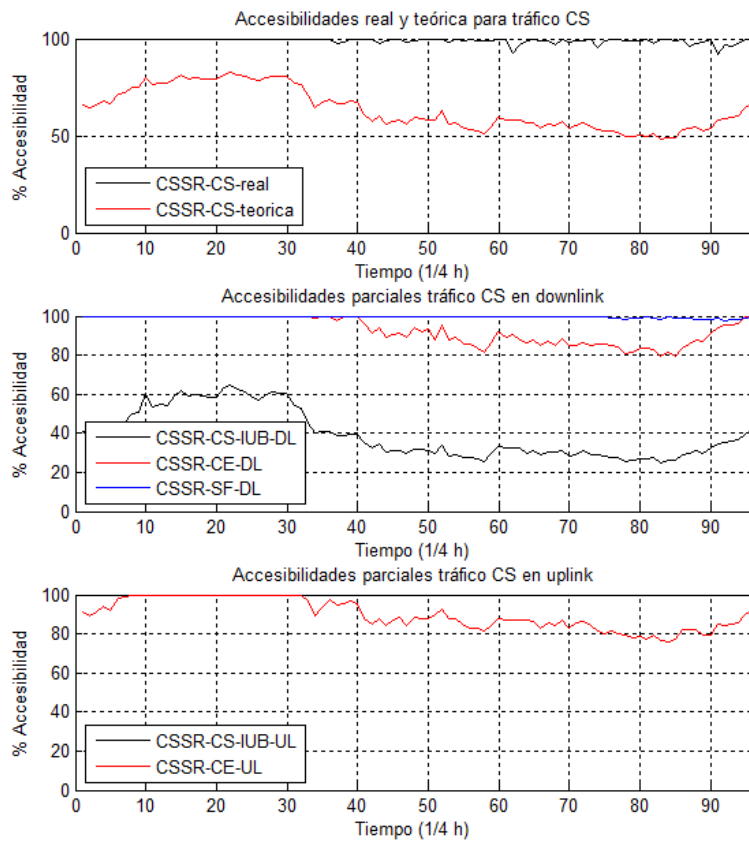


Figura 8.41: Accesibilidades real y teórica para el tráfico CS en el nodo C (después)

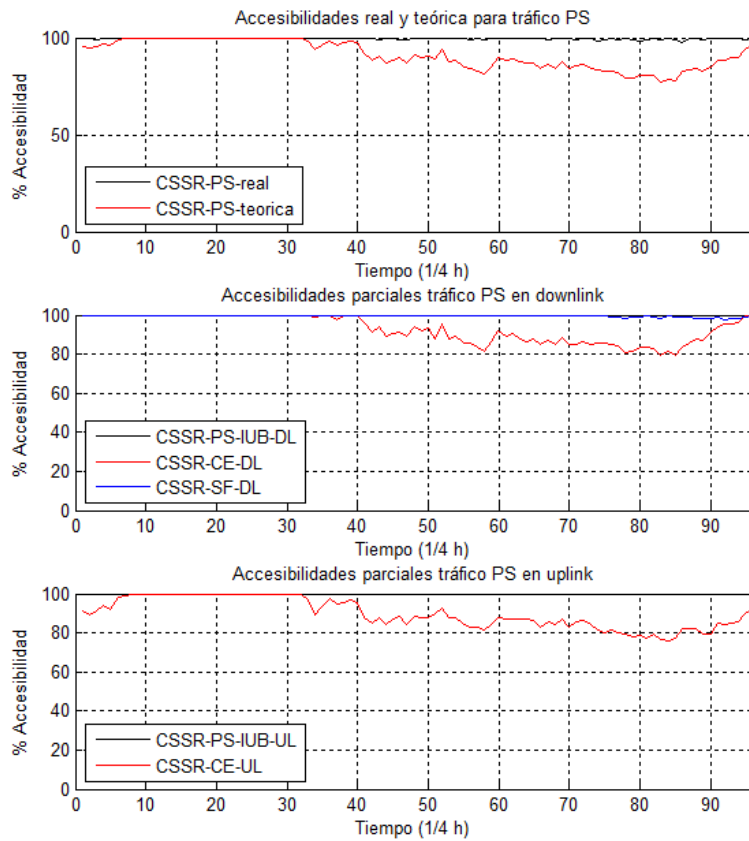


Figura 8.42: Accesibilidades real y teórica para el tráfico PS en el nodo C (después)

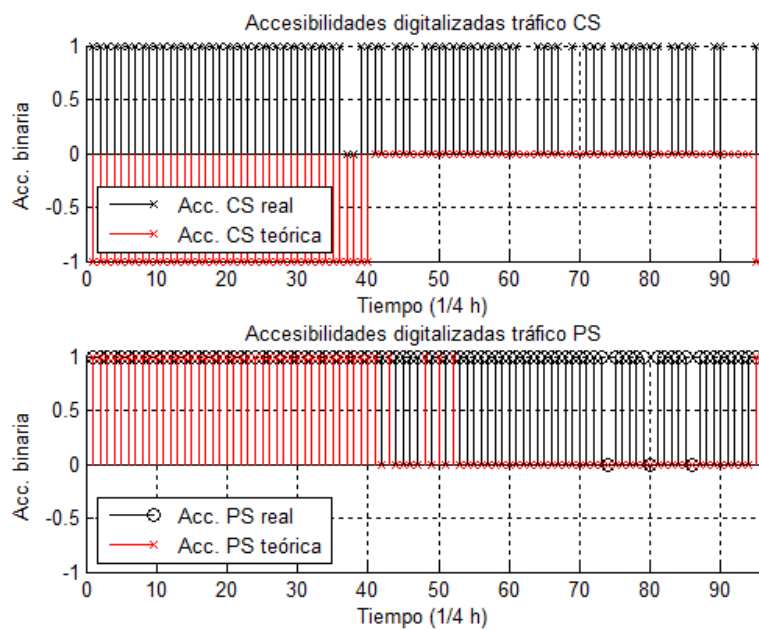


Figura 8.43: Predicción de la degradación en modelos teórico y real (nodo C, después)

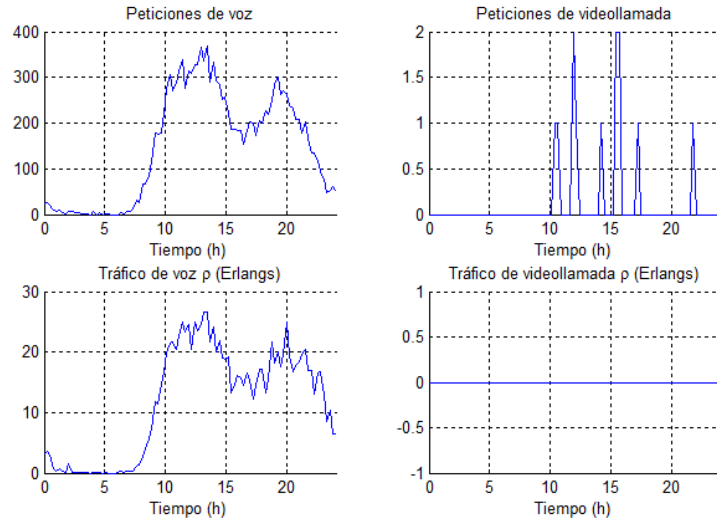


Figura 8.44: Tráfico CS durante un día de observación (nodo D, antes)

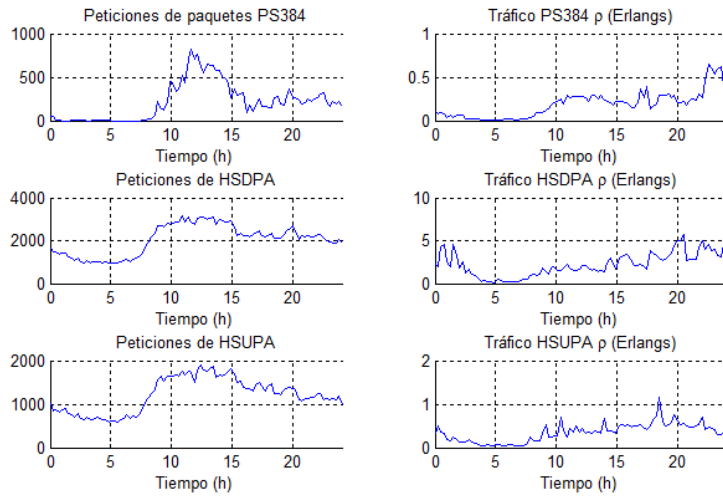


Figura 8.45: Tráfico PS durante un día de observación (nodo D, antes)

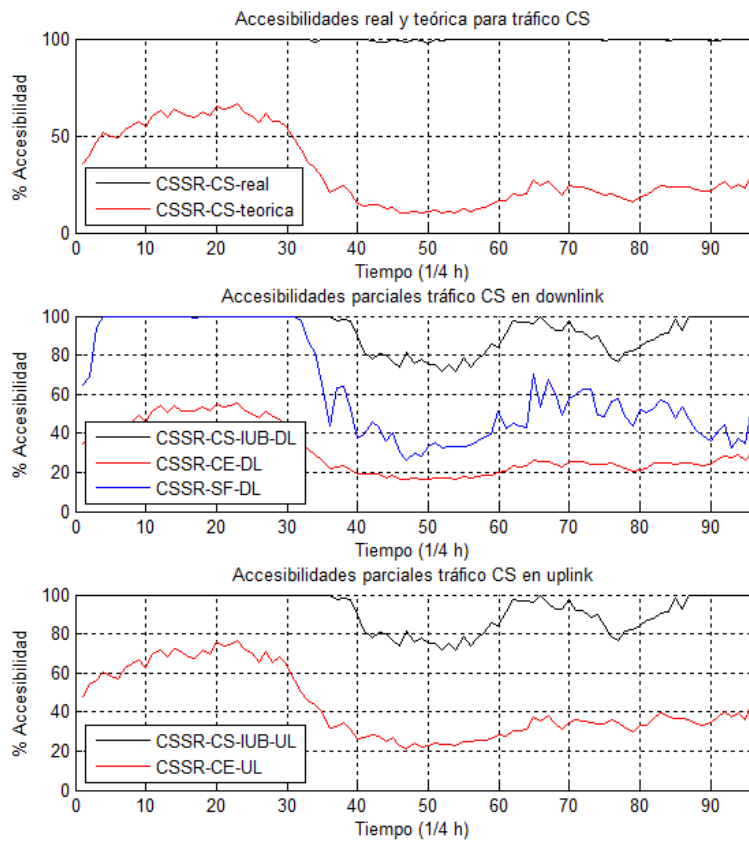


Figura 8.46: Accesibilidades real y teórica para el tráfico CS en el nodo D (antes)

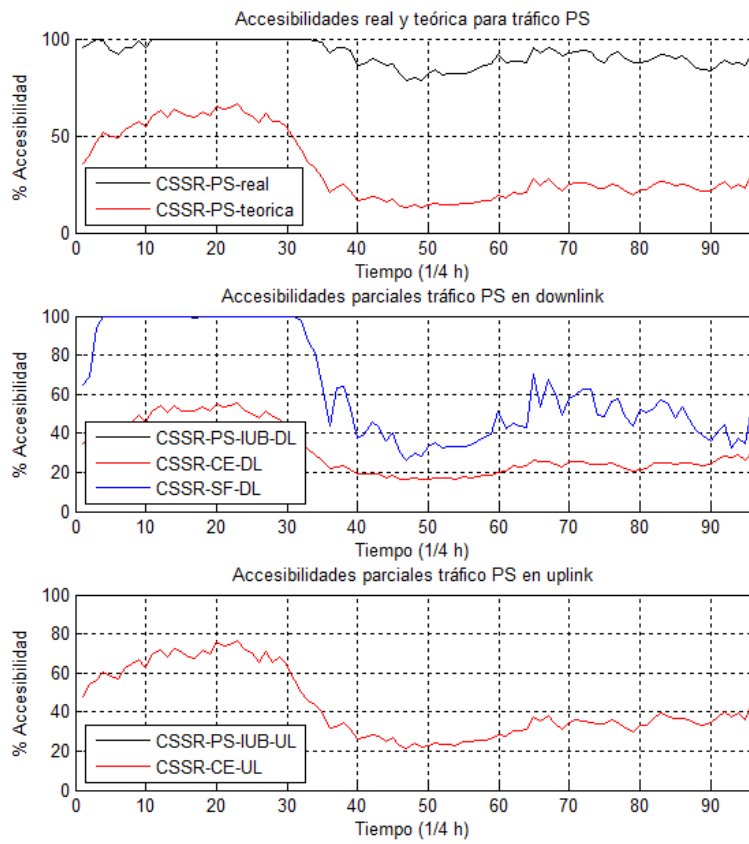


Figura 8.47: Accesibilidades real y teórica para el tráfico PS en el nodo D (antes)

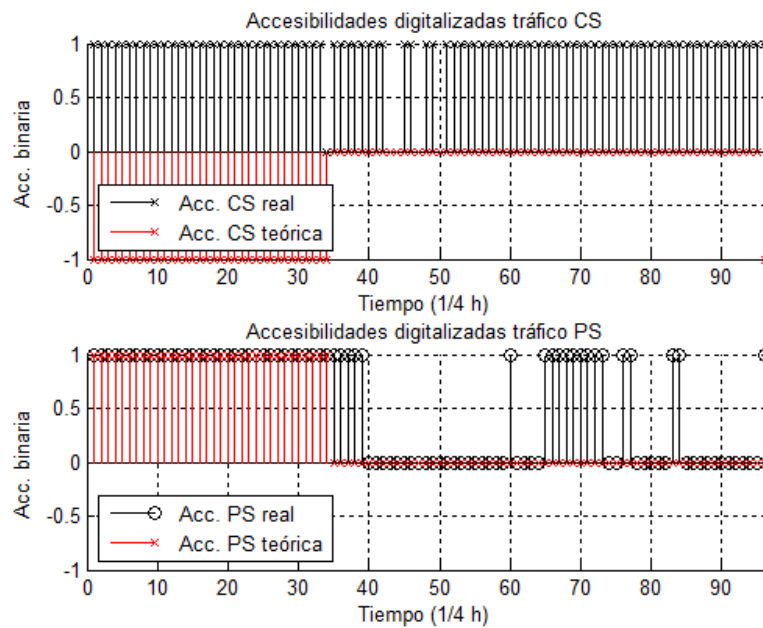


Figura 8.48: Predicción de la degradación en modelos teórico y real (nodo D, antes)

que la accesibilidad por Iub es notablemente mejor que los otros factores. Esto es debido a que en este caso hay una mayor capacidad del Iub de circuitos al ser un nodo Full-IP y realizar todo el transporte mediante ADSL y por ello es menos limitante. A la vista de las accesibilidades totales, se concluye que este nodo sufre un problema de congestión.

En la figura 8.48 se muestran las accesibilidades digitalizadas.

8.4.4.2. Escenario 8

Para evitar que haya tanta congestión en este nodo se opta también por aumentar el número de CE (figura 8.49). De este modo, en este caso, se aumentan el número de tarjetas en el enlace ascendente.

⇒ CE UL: 1152 CE.

⇒ CE DL: 768 CE.

⇒ SC: 1 portadora.

⇒ Iub UL: 30 Mbps.

⇒ Iub DL: 10 Mbps.

En las figuras 8.50 y 8.51 se recogen los tráficos de entrada para el modelo.

Al aumentar el número de tarjetas de CE_UL se aumenta el número de Channel Elements en el enlace ascendente, con lo cual la accesibilidad mejora respecto a lo visto en el escenario anterior, tal y como demuestran las figuras 8.52 y 8.53.

Por último, en la figura 8.54 se representa la predicción de la degradación que resulta de la digitalización de la accesibilidad teórica y real.

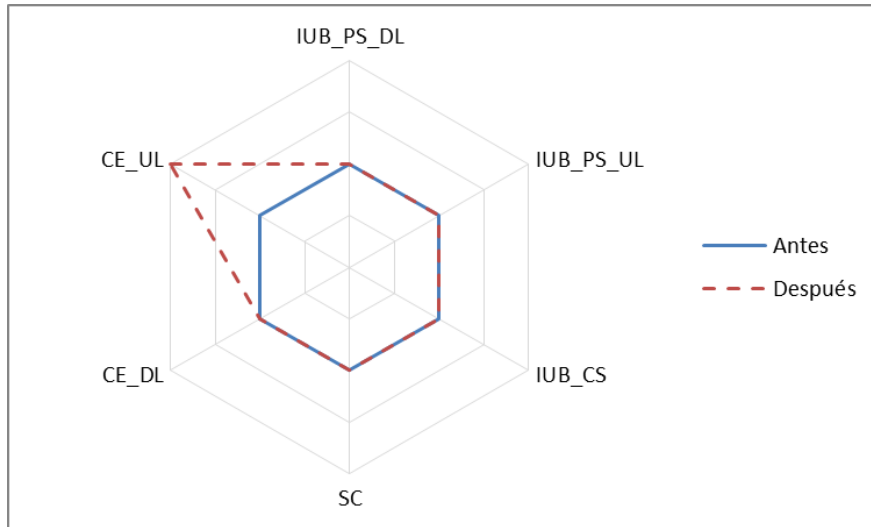


Figura 8.49: Ampliación de recursos en el nodo «D»

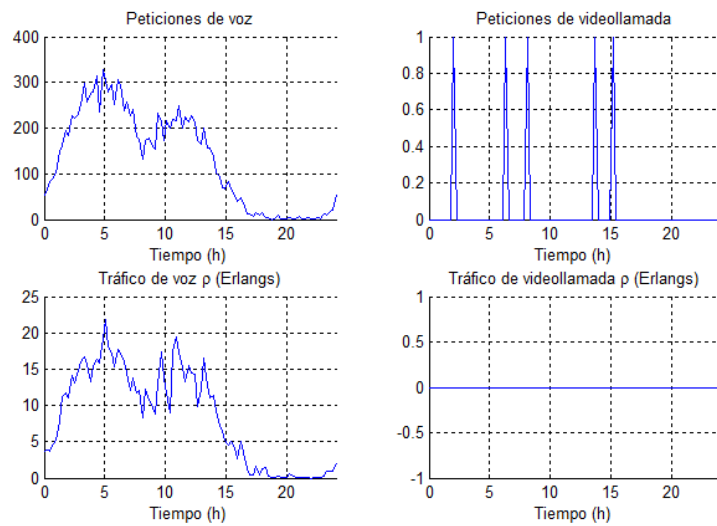


Figura 8.50: Tráfico CS durante un día de observación (nodo D, después)

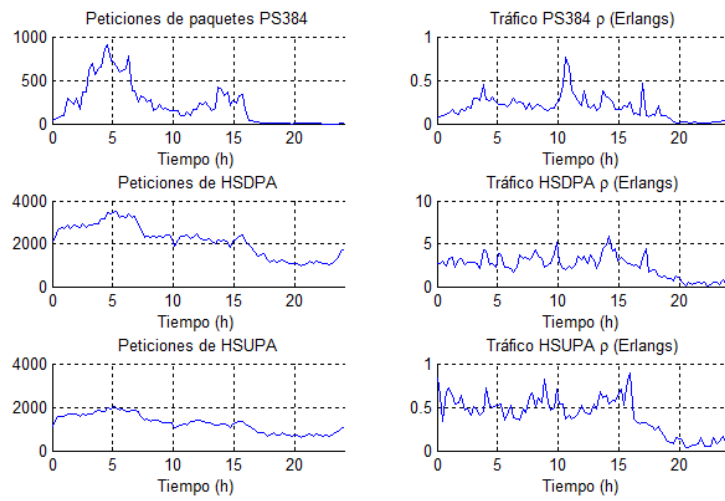


Figura 8.51: Tráfico PS durante un día de observación (nodo D, después)

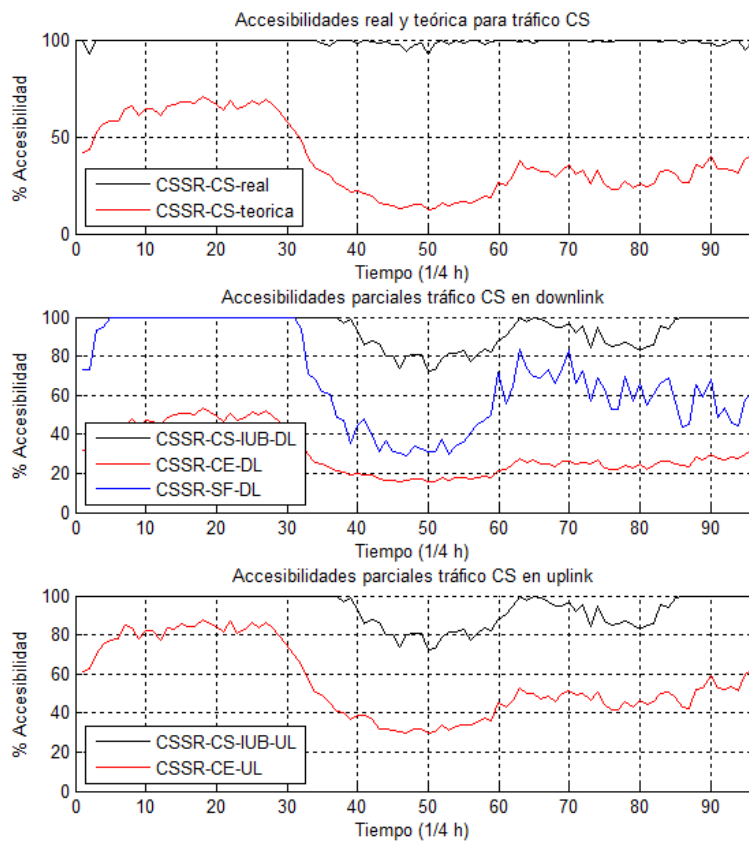


Figura 8.52: Accesibilidades real y teórica para el tráfico CS en el nodo D (después)

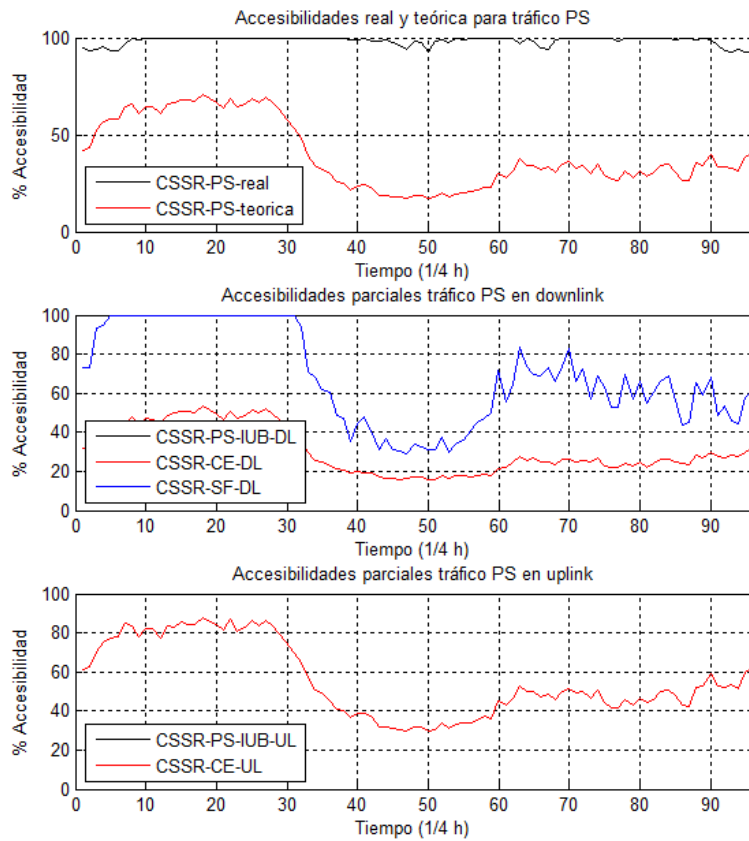


Figura 8.53: Accesibilidades real y teórica para el tráfico PS en el nodo D (después)

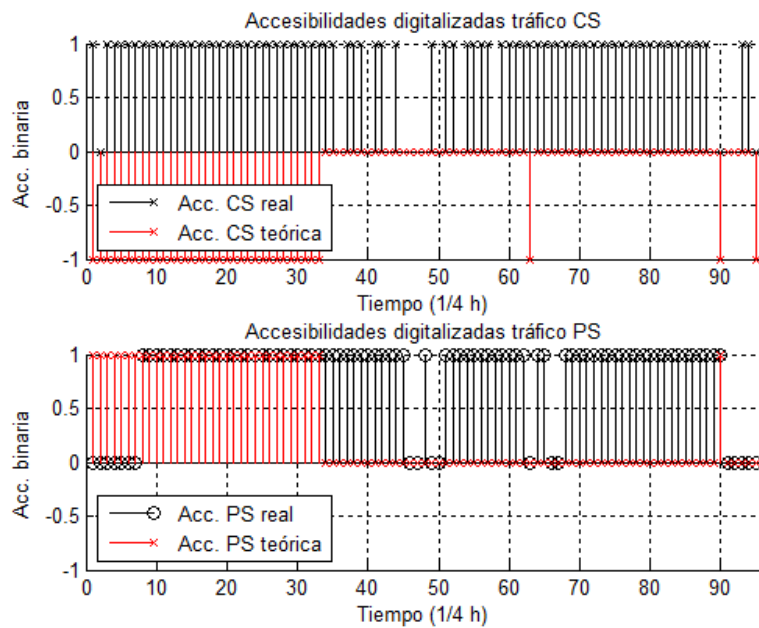


Figura 8.54: Predicción de la degradación en modelos teórico y real (nodo D, antes)

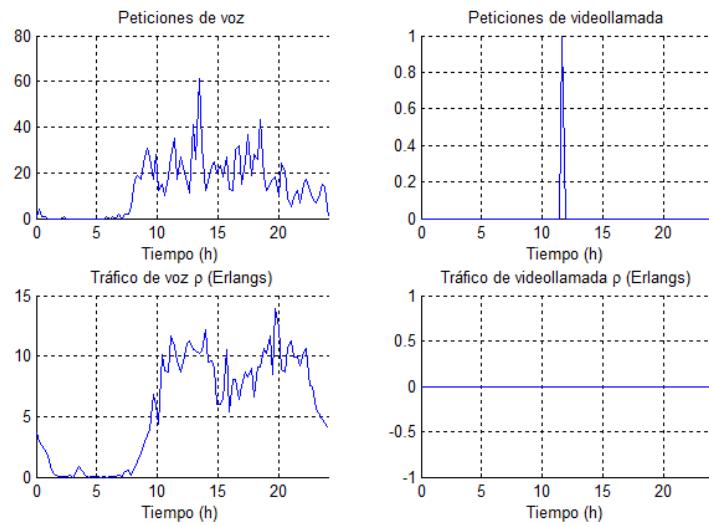


Figura 8.55: Tráfico CS durante un día de observación (nodo con fallos HW)

8.4.5. Nodo con fallo hardware

Otra de las contribuciones que aporta el trabajo de esta Tesis es la capacidad de poder distinguir en un nodo una degradación por congestión de un fallo hardware. Se ha estudiado un nodo real sobre el que el operador conocía que tenía problemas hardware. Este nodo está equipado de la siguiente manera:

⇒ IUB CS = 1.272 Mbps.

⇒ IUB PS UL = 4.096 Mbps.

⇒ IUB PS DL = 30 Mbps.

⇒ CE UL = 384 CE.

⇒ CE DL = 768 CE.

⇒ SC = 1 árbol (1 portadora).

Se introducen estos parámetros, así como el tráfico cursado (figuras 8.55 y 8.56), en el modelo. Los resultados en cuanto a accesibilidad para tráficos CS y PS se muestran en las figuras 8.57 y 8.58, respectivamente.

Teniendo en cuenta que los fallos de conexión se definen como:

$$Bloqueo_{total} = Bloqueo_{congestión} + Bloqueo_{HW} \quad (8.5)$$

puede darse la situación de que el bloqueo por congestión (que es el que predice el modelo) alcance un determinado valor, aún cuando los datos reportados por el nodo sugieren que la accesibilidad sea todavía menor. En ese caso se puede concluir que los fallos origen de esa degradación no son por congestión sino por fallo hardware.

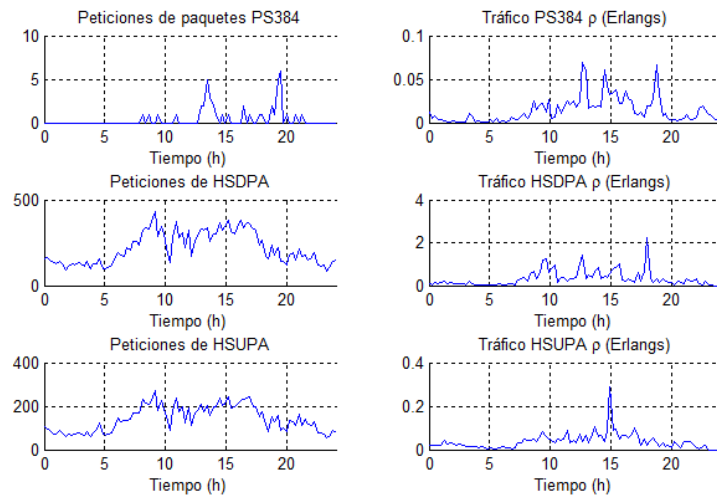


Figura 8.56: Tráfico PS durante un día de observación (nodo con fallos HW)

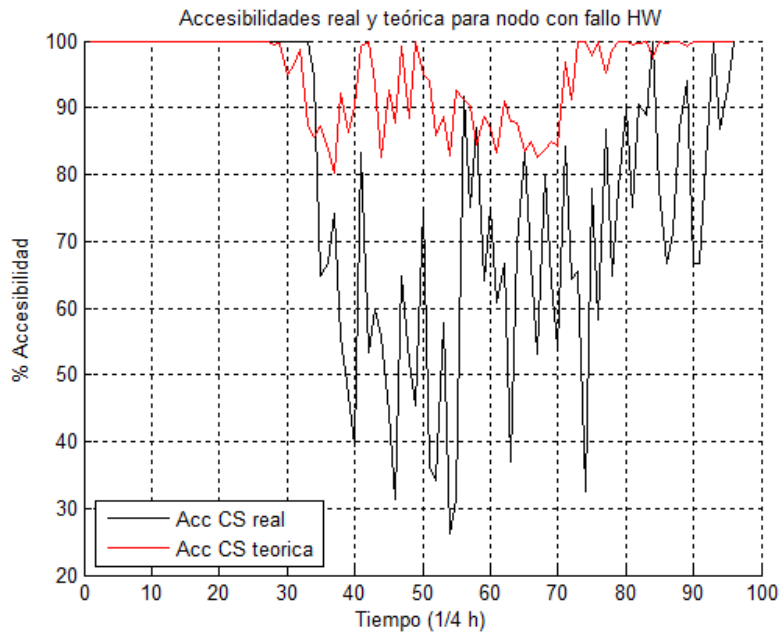


Figura 8.57: Accesibilidades real y teórica para tráfico CS en nodo con fallo hardware

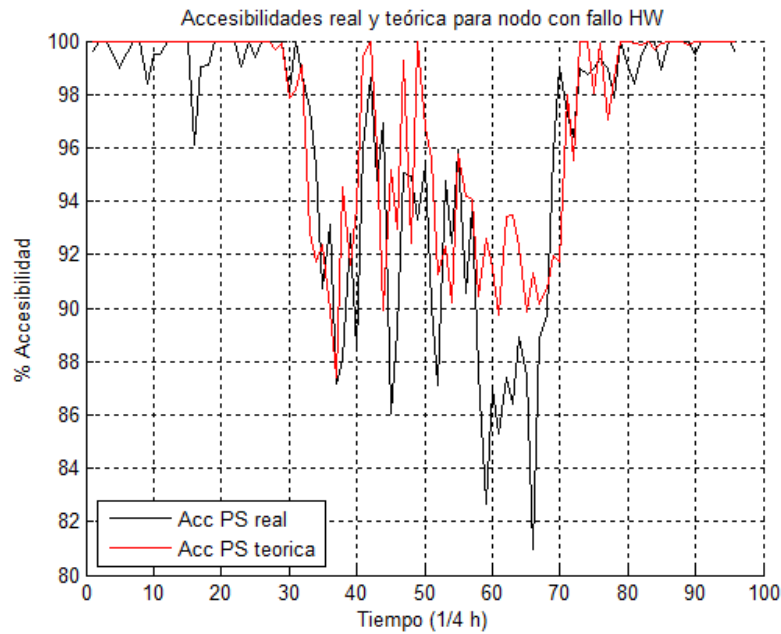


Figura 8.58: Accesibilidades real y teórica para tráfico PS en nodo con fallo hardware

8.5. Conclusiones

El objeto del presente trabajo consiste en modelar el consumo de capacidad ejercido por los diferentes servicios que atiende el nodo UTRAN. Sólo se tendrán en cuenta los recursos susceptibles de ser ampliados por medio de operación de red (tarjetas de procesado banda base, portadoras WCDMA con su árbol de códigos y capacidad Iub). Como ya se ha visto, se modela la accesibilidad para cada uno de estos recursos, y posteriormente se estima la accesibilidad agregada. A sabiendas de que no se están teniendo en cuenta otros recursos que impactan en la accesibilidad agregada (como la interferencia uplink debida al solape entre celdas y a la carga que generan los usuarios), la potencia downlink o la capacidad en la RNC, se comparan las accesibilidades teórica y la real reportada por los Key Performance Indicators de la red. En esta comparativa, se aprecia que hay diferencias en los valores absolutos, como es evidente al dejar fuera del modelo parte de los recursos implicados, pero sí se puede constatar que las variaciones de tendencia son análogas en el modelado teórico y en la realidad. De esta forma es posible detectar las degradaciones de accesibilidad aunque no se consiga dar el valor exacto de ésta. La ventaja es que es posible predecir cuándo un nodo se va a degradar, y además, se podrá saber qué recurso es el limitante, para poderlo ampliar. Los KPI reportados actualmente por la red no están desglosados por recurso, con lo que no es posible conocer cuál es el limitante. Además, se podrán diferenciar fallos de conexión asociados a congestión de fallos debidos a mal funcionamiento del hardware.

Capítulo 9

Conclusiones y futuras líneas de trabajo

9.1. Conclusiones

El marco de estudio de la presente Tesis se ha centrado en la red de acceso radio para redes de tercera generación (UTRAN). En este contexto, se han estudiado los diferentes tipos de servicios que soporta dicha red, distinguiendo entre servicios rígidos (voz, videollamada y su señalización asociada) y elásticos (tráfico de paquetes PS384 y HSPA). Se ha caracterizado también la relación existente entre servicios y los recursos que consumen. Se tiene un modelo en el que el número de recursos es limitado, vulnerable a problemas de congestión, y todo esto en un contexto en el que el despliegue de red es inevitable para dotar a los usuarios finales de nuevos servicios.

Dos son las líneas de trabajo de las que se pueden aprovechar los resultados de la presente Tesis:

- ⇒ Por una parte, en una situación en la que se suceden anomalías es posible a posteriori parametrizar este contexto y, basándose en un modelo teórico, predecir si se trata de un fallo hardware o si es un problema de congestión. Al mismo tiempo, la Tesis se anticipa al momento en que la tecnología actual pueda proporcionar estadísticos con mayor desglose, lo que permitiría aplicando las propuestas planteadas determinar cuál es el recurso limitante en la red.
- ⇒ Bajo otro enfoque, el disponer de un modelo teórico válido y fiable, permite anticipar a priori posibles problemas de congestión. Debe recordarse en este punto que el dimensionamiento de una red de telefonía es un proceso costoso, y que actualmente los operadores de red no cuentan con herramientas específicas para determinar recurso a recurso las consecuencias de la ampliación. Por el contrario, su metodología consiste en implantar los equipos en un emplazamiento y plantear después las acciones correctivas que sean necesarias. Partiendo de los mismos supuestos enunciados en el punto anterior, sería posible anticiparse y determinar cuál es el recurso en el que se debe invertir con mayor prioridad para garantizar un buen funcionamiento.

El modelo propuesto permite estimar la probabilidad de bloqueo en una red UTRAN conocida la demanda para cada uno de los servicios y el conjunto de recursos disponible. Dado que en una red UTRAN es conocido el % de bloqueo real, puede deducirse si este bloqueo se debe a

congestión en el sistema o a fallo hardware. La congestión podrá subsanarse ampliando recursos, y el fallo hardware conllevará tareas de operación y mantenimiento. En el caso de necesitar una ampliación de capacidad, el modelo también permite determinar cuál es el recurso más limitante. La descongestión será mayor si se amplía el recurso que más limita.

9.1.1. Beneficios del uso de este modelado en la operación de redes UTRAN

En los inicios de la Tesis se ha hecho un estudio exhaustivo del estado del arte en los términos de garantía de calidad de servicio por parte de las redes UTRAN y de la forma de trabajo de las operadoras de red móvil. Como conclusiones de este estudio previo se pueden extraer:

- ⇒ A día de hoy, existen estudios que modelan los procesos de control de admisión (CAC) para la garantía de QoS bajo diferentes puntos de vista, siempre desde un enfoque teórico y académico, pero sin contrastar ni validar sus modelos con datos reales. Estas investigaciones están basadas en estudios y simulaciones del entorno radio en UMTS. En definitiva, sólo están basadas en simulaciones en el momento del despliegue. Como ya se ha comentado en el capítulo de introducción, una de las motivaciones de la presente Tesis consiste en construir un modelo fiable y válido que reúna en un mismo plano los estudios teóricos con datos reales reportados por los nodos. Esta Tesis desarrolla un modelo enmarcado en el ámbito de una red ya desplegada, que tiene en cuenta los factores que verdaderamente un operador de red es capaz de variar. Gracias a este trabajo, un operador de red sabe qué recurso tiene que dimensionar en cada momento.
- ⇒ Por otra parte, la forma de trabajo de los operadores de red prioriza en muchos casos un despliegue rápido sin una planificación exhaustiva de los recursos, sobredimensionando en muchos casos algunos, dejando en segundo plano otros más limitantes. Esta forma de trabajo, con una filosofía más correctiva, hace a los operadores fuertemente dependientes de datos estadísticos y medidas, y al mismo tiempo se trata de un proceso costoso por la inversión tecnológica que es preciso llevar a cabo para la compra e instalación del equipamiento. Ninguno de los operadores utiliza herramientas ni modelos que permitan priorizar la inversión en recursos.

9.1.2. Modelado multiservicio en los subsistemas UTRAN

La heterogeneidad característica de los tráficoes que soporta la red UMTS plantea un escenario en el que se tienen tráficoes que en principio demandan individualmente diferente cantidad de recursos. Además, a diferencia de las redes GSM, en las redes UTRAN hay diferentes recursos cuya capacidad es gestionada de forma independiente: capacidad de procesado banda base, capacidad de spreading code WCDMA disponibles en las portadoras radio y capacidad en la interfaz Iub, distinguiéndose flujos garantizados (como los ATM a tasa binaria constante y tiempo real o los MPLS CRB RT – usados en servicios de voz), de los no garantizados (como los ATM o MPLS VBR NRT -empleados en servicios de datos).

El sistema, dispone de una cantidad de recursos limitados. Por tanto, todos los tráficoes «compiten» al mismo tiempo por un mismo recurso.

Esta situación ha llevado al autor a pensar en la aplicación directa de un algoritmo que permita analizar el bloqueo de cada uno de los servicios. La implementación de este algoritmo y su

posterior validación numérica han dado pie a que durante la Tesis se desarrolle una investigación más exhaustiva en este campo. Las conclusiones que al autor de esta Tesis plantea en relación con este apartado son las siguientes:

- ⇒ La naturaleza de la red UTRAN como red mutiservicio hace evidente que no es posible utilizar el modelo de Erlang-B para el modelado de la misma. No obstante, un modelo de análisis de bloqueo como el modelo de Kaufman-Roberts se ajusta de una forma muy precisa a la situación de bloqueo multiservicio característica de una red UTRAN, como así lo atestigua el ensayo numérico de diferentes escenarios reales y su contraste con las estadísticas reales.
- ⇒ El proceso en su conjunto requiere, para su validación, la ejecución iterativa del mismo algoritmo abstracto, particularizado a diferentes situaciones (tipos de tráfico rígido, elástico, enlace ascendente y descendente). Por tanto, es imprescindible encontrar un algoritmo fiable y con una complejidad algorítmica baja para que la herramienta sea manejable. El algoritmo primitivo planteado por Kaufman y Roberts en la década de los 80 funciona muy bien cuando se trabaja con un número de peticiones pequeña y con una capacidad de sistema limitada. La validación del algoritmo bajo cantidades de mayor magnitud da como resultado el desbordamiento numérico y, dada su alta complejidad algorítmica, una ejecución muy lenta, por lo que en estos supuestos no es recomendable ni adecuada la implementación de este algoritmo.
- ⇒ Existen otros planteamientos del algoritmo de Kaufman-Roberts mejorados que paliar con éxito las carencias del algoritmo original. En particular, y como base para el desarrollo de la Tesis, se han consultado diversas fuentes y se ha implementado una sucesión de algoritmos que han ido mejorando la calidad de las simulaciones. Un algoritmo concreto basado en la FFT ha resultado ser válido y fiable, evita el desbordamiento numérico y al mismo tiempo posee una complejidad algorítmica mucho más sencilla, lo que se traduce en cálculos más rápidos.

9.1.3. Modelado y conformación del tráfico UTRAN

En capítulos previos se ha hecho patente que el diseño del modelo conlleva la necesidad de modelar adecuadamente el tráfico de la red UTRAN, utilizando las variables del conformado como parte fundamental del proceso de calibración del sistema. A este respecto, se pueden concluir los siguientes puntos:

- ⇒ A día de hoy, las estadísticas de tráfico (y en general para otras medidas también) reportadas por los Nodos B tienen una utilidad limitada, desde el punto de vista de que una resolución de 4 medidas agregadas por hora supone una información muy pobre cuando lo que se pretende es estudiar un problema de congestión, ya que la duración del servicio de manera individual para cada cliente puede ser mucho menor. Diversas simulaciones realizadas durante el desarrollo de la Tesis dejan patente que esa resolución temporal desvirtúa completamente los resultados, haciendo «invisible» la verdadera degradación del sistema. Por ello, el autor concluye que es conveniente disminuir la resolución temporal, estableciendo que una buena unidad de medida son los minutos. Con este requisito, el modelado del sistema debe interpolar el tráfico en la dimensión tiempo recurriendo a teoría estadística. Esta práctica trae consigo la aparición de diferentes variables con las que se trabaja durante el proceso de calibración del sistema.
- ⇒ La interpolación de tráfico está, como se ha comentado anteriormente, sujeta a la teoría matemática subyacente al planteamiento estadístico que se le ha dado a este problema a

lo largo de la Tesis. En primer lugar el autor ha caracterizado cómo debería distribuirse el estadístico agregado en los 15 minutos previos, concluyendo que la distribución no puede ser uniforme, ya que no modela una situación real, sino que por el contrario debe tener forma «acampanada». Además, la «campana» debe tener una media y varianza moderadas, ya que el hecho de inventanar la curva en un intervalo finito da como resultado un error de reconstrucción posterior. Haciendo una simulación específica para la curva de densidad de probabilidad de una distribución de Poisson se concluye que para valores de λ crecientes aumenta de forma exponencial el error de reconstrucción. Además, la muestra de tráfico reconstruida da como resultado patrones de tráfico de menor intensidad que corroboran la dependencia del tráfico en el modelado y cómo una mala elección da como resultado un problema desvirtuado.

⇒ Dadas las conclusiones extraídas del apartado anterior, se ha realizado un estudio de las distribuciones Normal y de Poisson. El análisis de cada una de ellas de forma independiente se ha realizado comparando las accesibilidades real y teórica de nodos de red de un operador real, desde un punto de vista cuantitativo, y de acuerdo a dos métricas: error cuadrático medio y coeficiente de correlación de Pearson. Asimismo, y para cada distribución en particular, se aprecia cómo la exactitud del modelo tiene una dependencia significativa de los parámetros de la distribución (media y varianza). Por el contrario, comparando ambas distribuciones no se aprecian diferencias significativas, por lo que el autor concluye que resulta más conveniente elegir una distribución de Poisson. Así, el modelo se lleva al mismo contexto que los procesos de Poisson estudiados en la bibliografía de teoría de colas, y además se tiene una sola variable λ en contraposición con la distribución Normal, donde el problema añade una dimensión al trabajar con el binomio (μ, σ) , llevando al diseño a un proceso de mayor complejidad innecesario.

9.1.4. Caracterización de los recursos

Una vez que ha quedado clara la intención de modelar el sistema utilizando una variante del algoritmo de Kaufman-Roberts, el objetivo inmediato es conseguir abstraer el problema real parametrizando y cuantificando las variables que intervienen en dichos algoritmos. Como conclusiones a este proceso, cabe citar:

⇒ Dado que el funcionamiento interno del sistema real es conocido, se puede apreciar con claridad que las diferentes etapas de procesado de las señales pueden modelarse como una cadena de subsistemas totalmente independientes. Es decir, que por ejemplo el bloqueo debido al procesado en banda base en el nodo B no está relacionado con la falta de disponibilidad de códigos de canalización. Puede hablarse de cinco subsistemas o cinco factores que son susceptibles de intervenir en el estudio: el procesado en banda base (cuyo recurso limitante son los Channel Elements), los códigos de canalización (únicamente en el downlink), la capacidad del enlace Iub, la potencia del amplificador en el downlink y las interferencias en el enlace ascendente. Cada uno de estos subsistemas puede abstraerse en el modelo de Kaufman-Roberts, obteniendo así sendas probabilidades de bloqueo parciales que deberán combinarse posteriormente.

⇒ En cualquier caso, este modelo sólo tiene en cuenta los tres primeros, obviando la implementación de los dos últimos (interferencia en el uplink y potencia en el downlink). El primero

está basado en el solape entre nodos y el segundo en la distancia a la que se encuentran los usuarios. Son factores que no dependen de la cantidad de tráfico de cada servicio, sino de la red. Este estudio se particulariza para los recursos que serán susceptibles de ampliación durante la operación de la red. Será posible incluso ampliarlos de forma proactiva una vez conocida la previsión de tráfico para determinados eventos en la red (los KPI reportados actualmente por la red no están desglosados por recurso, con lo que no es posible conocer cuál es el limitante).

9.1.5. Implementación del modelado

Resta decir que para el desarrollo de la Tesis resulta fundamental la simulación matemática del modelo bajo diferentes escenarios. Esto lleva al autor a tomar una decisión sobre cuál es la herramienta más adecuada para implementar los algoritmos necesarios. Como conclusiones a este respecto cabe citar:

- ⇒ Como requisitos para la implementación de un modelo de este tipo es necesario contar con una herramienta de análisis matemático robusta, dados los requisitos que conlleva la ingente cantidad de datos que es necesario manejar, sin perder de vista la complejidad algorítmica que demanda. Se concluye que los cálculos deberán efectuarse de la forma más abstracta posible, encapsulando en la medida de lo posible las funciones más básicas de los subsistemas, lo cual facilitaría la comprensión del modelo y permitiría implementar cambios sustanciales de forma eficaz.
- ⇒ Existen multitud de herramientas en el mercado a las que se puede recurrir para implementar el modelo. El autor ha barajado algunas como Opnet Modeler, Visual Basic, C#, Java, Python y Matlab. Después de valorar las diferentes alternativas, el autor concluye que la más conveniente para el modelado es Matlab. Dado que el trabajo de la Tesis está basado en un modelo matemático, ésta resulta ser la herramienta más adecuada, ya que implementa de serie gran cantidad de funciones matemáticas y representaciones gráficas que permiten ahorrar tiempo en la programación del sistema. Al mismo tiempo, ofrece un lenguaje de programación propio (lenguaje M) que permite encapsular las funciones definidas por el autor, de acuerdo con los requisitos establecidos en el punto anterior.

9.1.6. Calibración del sistema vs. modelado del tráfico

Como se ha comentado anteriormente una de las principales aportaciones de la Tesis consiste en aprovechar el conocimiento que el autor tiene del funcionamiento real de los equipos, además de la disposición de los datos reportados por nodos de un operador de red real, para la elaboración y validación de un modelo teórico construido expresamente para esta Tesis. La calibración del sistema se beneficia pues de esta información. Tras el proceso de diseño y validación de esta etapa del proceso, las conclusiones del autor son las siguientes:

- ⇒ Es imprescindible tener presente en todo momento la idea de que el éxito del diseño del modelo debe ser cuantificado en términos de calidad, definiendo una política razonable antes de la implementación del modelo. Así, el autor hace patente que el mejor modelo no es el «mejor» ni el «menos malo» en comparación con otras alternativas, sino que el modelado estará calibrado cuando el sistema cumpla unos objetivos mínimos de calidad. Como medida

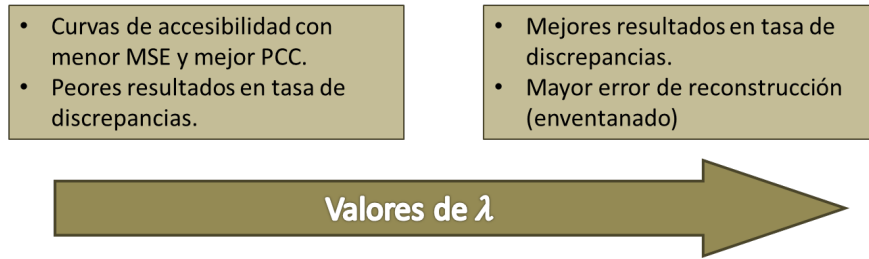


Figura 9.1: Compromiso en la elección del parámetro λ para una distribución de Poisson

de esta calidad, se plantea en un primer momento la necesidad de que el sistema discrimine entre fallos hardware y problemas por congestión.

- ⇒ Esta degradación no es necesario que se mida en términos cuantitativos, en el sentido de que el modelo no debe pretender «seguir» una curva. No se trata pues de obtener a la salida una curva evolutiva de accesibilidad igual a la real, sino simplemente de detectar cuándo se produce la degradación. Por ello, el autor concluye que es necesario enfocar el problema desde el punto de vista binario o digital.
- ⇒ Derivado del punto anterior, las señales serán discretizadas en base a un umbral. Teniendo presente que las curvas teóricas presentan un mayor rizado que las reales, el autor propone que una metodología para minimizar la tasa de discrepancias entre el modelo teórico y el sistema real sería recurriendo a la calibración de un umbral de discretización para la accesibilidad teórica personalizado para cada caso de uso. En este punto, se han realizado simulaciones con diferentes escenarios, para diferentes perfiles de tráfico y con diferentes umbrales teóricos para cuantificar su impacto en la calidad final del sistema. La conclusión que se puede extraer del experimento es que la calidad final del sistema no se ve afectada por la distribución de tráfico sea cual sea el escenario, sino que por el contrario existe un umbral teórico estable invariante con el tráfico y con el escenario que minimiza la tasa de discrepancias.
- ⇒ La elección de los parámetros de la distribución con la que se interpola el tráfico supone una solución de compromiso entre diversos factores. En la figura 9.1 se refleja la elección del parámetro λ para una distribución de Poisson. Valores más pequeños de λ dan como resultado curvas de accesibilidad en función del tiempo con formas similares. A nivel general y comparando cuantitativamente la accesibilidad real con la teórica calculada de esta forma, el error cuadrático medio y el coeficiente de correlación de Pearson presentan mejores resultados con valores pequeños. No obstante, si se pretende minimizar la tasa de discrepancias en el pronóstico de la degradación, debe elegirse un valor de λ mayor, teniendo presente que a medida que éste crece aumenta el error de reconstrucción de la interpolación previa del tráfico.

9.2. Líneas futuras

A lo largo de la presente Tesis se ha elaborado un modelo teórico para redes multiservicio con tecnología UMTS, y se ha validado recurriendo al conocimiento que tiene el autor sobre el funcionamiento real de los equipos, así como a estadísticas reales reportadas por los nodos en

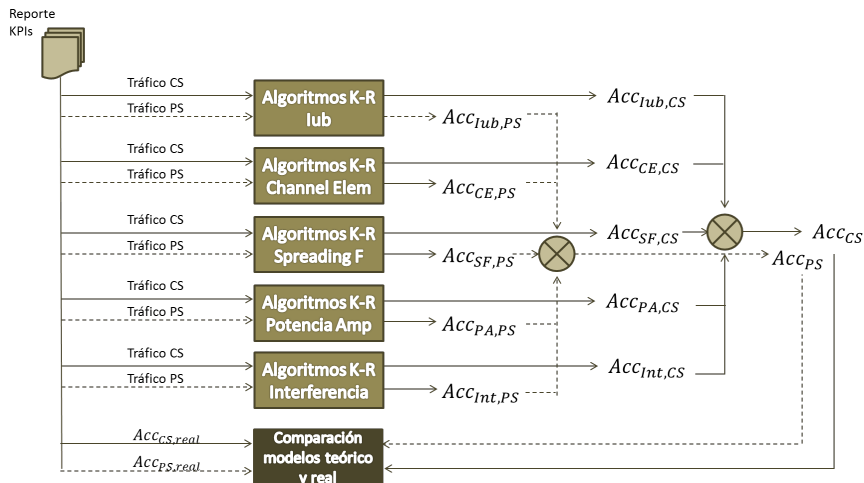


Figura 9.2: Aplicación multietapa del algoritmo de K-R en el *downlink* ampliado

diferentes escenarios. El objetivo de este apartado es esbozar algunas ideas sobre investigaciones futuras que pueden tomar como base los resultados de esta Tesis.

9.2.1. Caracterización de servicios

En la documentación de la Tesis se ha dedicado un capítulo expresamente a la caracterización de servicios en las redes UTRAN. Así, se han estudiado los servicios portadores y las diferentes clases de tráfico que definen las políticas de QoS en tecnología UMTS. Particularmente, en el caso de los smartphones, en la actualidad se está poniendo de manifiesto el elevado tráfico de señalización (plano de control) que éstos generan. Esta señalización puede llegar a ser tan significativa que provoque congestión en la red, por lo que se propone la caracterización de estas conexiones y el estudio del impacto que este tipo de tráfico puede tener en la degradación de la QoS.

9.2.2. Consumo de recursos

Como ya se ha comentado anteriormente, durante el desarrollo del modelo se ha tratado de elaborar un modelo basado en un esquema modular, teniendo presente el funcionamiento interno del sistema real. No obstante, algunos de los subsistemas que se reflejan de manera teórica no pudieron ser tenidos en consideración, ya que el nodo no reporta de manera exhaustiva la información necesaria. Así, para poder cuantificar los efectos de la potencia del amplificador en el DL, o el margen de interferencia en el UL como recursos limitados en el sistema, es necesario conocer por ejemplo los usuarios que disfrutan de cada servicio. Se propone como futura línea de trabajo la extensión del modelo teniendo en cuenta estos recursos cuando tecnológicamente sea posible disponer del reporte de estos datos. Las figuras 9.2 y 9.3 muestran el contexto en el que podrían incluirse en un modelo mejorado.

Así mismo, existen otra serie de factores que tampoco se han estudiado, como los recursos de capacidad en la RNC, cuyo impacto en el modelo puede ser estudiado. Por último, otras tecnologías que se han ido introduciendo en el documento, como HSPA+ y LTE tampoco han sido modeladas, que sin duda constituyen una nueva línea de trabajo.

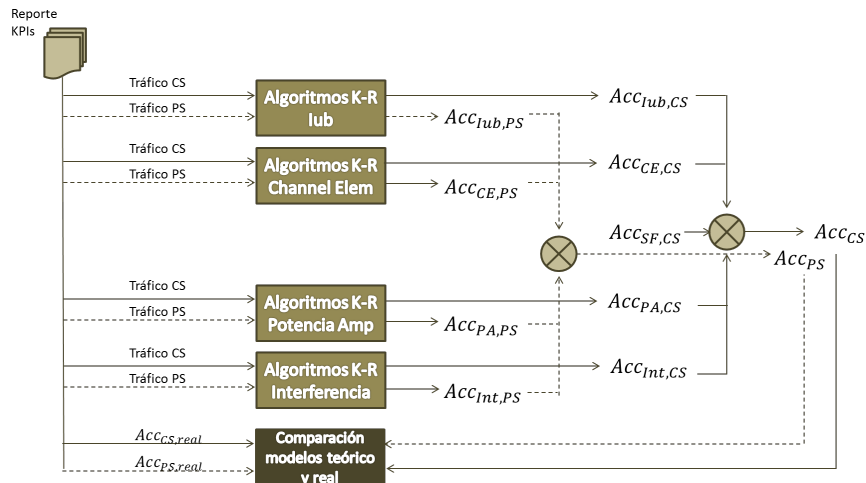


Figura 9.3: Aplicación multietapa del algoritmo de K-R en el *uplink* ampliado

9.2.3. Detección del factor limitante

El sistema que se ha diseñado hasta el momento permite, a partir de un modelado de tráfico y de los recursos disponibles, determinar cuándo una degradación en el sistema es debida a congestión por falta de recursos, descartando así cualquier fallo hardware.

El objetivo en este punto es, en aquellos casos en los que se concluya que existe degradación, determinar cuál es el factor causante de la degradación (factor limitante).

Actualmente, como se ha comentado en capítulos anteriores, en los KPIs reportados por el Nodo, figura como estadístico la accesibilidad real general, que como se ha visto es combinación de los enlaces ascendente y descendente. Sin embargo, la tecnología actual no es capaz de proporcionar estadísticas de cada subsistema, de manera que todas las accesibilidades teóricas parciales que se han calculado hasta el momento Acc_{iub} , Acc_{SF} , Acc_{CE} no tienen su homólogo real. Por ello, esta Tesis se limitará a explicar de manera descriptiva cómo se completaría el diseño con esta nueva prestación.

Modelado teórico generalizado

Como se ha descrito en capítulos anteriores, el sistema bajo estudio se ha modelado como una sucesión de subsistemas diferentes, cada uno de los cuales puede abstraerse con sus recursos y capacidad para aplicarles de manera independiente el algoritmo de Kaufman-Roberts. El bloqueo y la accesibilidad globales pueden ser entendidos como el producto de cada uno de estos subsistemas. Así pues, se ha presentado de manera latente un proceso que incluye 3 factores: la capacidad de procesado en banda base (recurso channel elements), los códigos de canalización y la transmisión (capacidad del enlace Iub). Al mismo tiempo, se han presentado otros dos factores, la potencia del amplificador y los niveles de interferencia en el uplink. Todo ello en un contexto en el que tecnológicamente el Nodo B no reporta la información de una manera lo suficientemente exhaustiva como para ser tenidos en cuenta en esta Tesis de manera cuantitativa.

De manera general, puede afirmarse que la accesibilidad asociada a un determinado Nodo B viene condicionada por un conjunto de factores, modelándose el conjunto como una cascada de respectivos subsistemas. La ecuación 9.1 representa el cálculo general de la accesibilidad teórica para un enlace y un tipo de tráfico concretos.

$$Acc_{teórica} = \prod_{j=1}^n Acc_j \quad (9.1)$$

Así, suponiendo que «n» factores condicionan el modelo, la accesibilidad teórica puede determinarse como el producto de cada una de las accesibilidades parciales asociadas (para el factor j-ésimo, accesibilidad Acc_j).

No obstante, el modelo anterior puede generalizarse aún más tratando de cuantificar la importancia o el peso que puede tener la degradación de un determinado elemento en su conjunto. La expresión 9.2 representa un modelo en el que la accesibilidad aparece como una ponderación de las accesibilidades parciales. Así, se introduce en el modelo un vector de coeficientes \vec{k} que pueden actuar como variables en una nueva calibración.

$$Acc_{teórica} = \prod_{j=1}^n k_j \cdot Acc_j \quad (9.2)$$

9.2.4. Degradación en el GoS

El modelo desarrollado en la Tesis es capaz de detectar la degradación del grado de servicio, diferenciando las anomalías debidas a fallos de tipo hardware de problemas de congestión. Se propone la profundización en el método de detección de la degradación del GoS, usando las siguientes técnicas:

- ⇒ El desarrollo de reconocimiento de patrones, que trabaja en base a un conjunto previamente establecido de todos los posibles objetos individuales a reconocer.
- ⇒ Las técnicas de agrupamiento o clustering, donde se realiza un modelado vectorial con diferentes características.
- ⇒ Por último, el empleo de redes neuronales, paradigma de aprendizaje y procesamiento automático, conforman una línea de trabajo alternativa, ya que si actualmente se recogen KPIs ya pasados, el desarrollo de una red neuronal permite la elaboración de predicciones.

9.2.5. Aplicación de técnicas de aprendizaje

Las técnicas de aprendizaje permiten modelar un sistema a partir del análisis de su comportamiento en un determinado periodo de observación. El modelo se irá evaluando con las nuevas observaciones, perfeccionando su caracterización matemática. Entre las técnicas más utilizadas se encuentran los árboles de clasificación, los algoritmos genéticos o las redes neuronales, que son descritas en el apéndice G. La aplicación al problema concreto puede tener la siguiente utilidad: a partir del modelo básico presentado en la tesis, se determina cuál es el recurso redundante, si ocurre un fallo hardware. Estas decisiones serán contrastadas con lo que luego se observe en la red, y así se irá corrigiendo el modelo, de forma que se mejore la tasa de acierto.

9.2.6. Implementación de la herramienta sw de modelado

Se propone como línea futura la implementación de estos modelos en una herramienta que pueda ser usada directamente por los operadores. Tendrá como entradas un fichero con la configuración del nodo y otro con las demandas de tráfico, y entregará a la salida la accesibilidad. También podrá fijarse la accesibilidad, y entonces entregará la cantidad de recursos necesarios en el nodo.

9.2.7. Aplicación a informes de estadísticas automáticos

Aplicación de estas técnicas en los informes de estadísticas. Usando los KPI reportados por la red relativos al tráfico, y la configuración de los nodos, se indicará para cada nodo si tiene déficit o exceso de recursos, o si tiene algún fallo hardware y en qué módulo.

9.2.8. Inclusión de otros factores limitantes de la red

El modelo podría incluir en el futuro otros factores limitantes, como el nivel de interferencia uplink (debida a la carga de usuarios y a la interferencia entre celdas), el consumo de potencia downlink, la distribución geográfica de los usuarios y la atenuación entre el nodo y cada usuario, la configuración de sistemas radiantes, y condiciones del entorno radio en términos de E_c/N_0 y de nivel de señal. También debería considerar la capacidad de la RNC, así como la parametrización radio: reparto de servicios por cada portadora, prioridad de cada servicio, etc.

Apéndice A

Código Fuente

A.1. Análisis de la accesibilidad en un escenario determinado

A.1.1. Rutina general

```
1 % FICHERO EJECUTABLE PARA EL ESTUDIO DE LA ACCESIBILIDAD CS Y PS
2 % =====
3 %
4 %   Notas:
5 %       - Pendiente digitalizar accesibilidades (¿Posible algoritmo para
6 %         detectar umbral de digitalización personalizado para cada curva?)
7
8
9 clear all, close all, clc
10
11 NumEscenario = 1;
12
13 % Estudio de la Accesibilidad Teórica CS y PS. Los indicadores que se
14 % obtienen tras la ejecución de este algoritmo son combinación de las
15 % accesibilidades parciales (cada factor IUB, CE, SF), cada una de ellas
16 % considerada en el DL y en el UL. Dichos valores Acc_CS_teorica y
17 % Acc_PS_teorica son métricas comparables a los KPIs reportados por el nodo
18 % B, CSSR_CS y CSSR_PS. Se devuelven también los mismos valores
19 % digitalizados en binario {1,0} indicando respectivamente "degradación" y
20 % "no degradación".
21 %
22 % Se devuelve también una matriz M que contiene, por filas, las
23 % accesibilidades parciales de la siguiente manera:
24 %
25 %   Fila 1: CSSR_CS_IUB_DL
26 %   Fila 2: CSSR_PS_IUB_DL
27 %   Fila 3: CSSR_CE_DL
28 %   Fila 4: CSSR_SF_DL
29 %   Fila 5: CSSR_CE_UL
30 %   Fila 6: CSSR_CS_IUB_UL
31 %   Fila 7: CSSR_PS_IUB_UL
32 %
33 % Como parámetro de entrada se necesita el nº de Escenario (NumEscenario).
34 % El resto de parámetros y estadísticas se cargan dentro de esta función.
35
```

```

[Acc_CS_teorica ,Acc_PS_teorica ,Acc_CS_digital ,Acc_PS_digital ,M]= estudio_acc(
    NumEscenario);
37
    % La función "digitalización" devuelve vectores de {1,0}. Los valores
39    % reales , los operadores consideran que hay degradación con menos del
    % 99%.
41    carga_estadisticas;
    CSSR_CS_digital = digitalizacion(CSSR_CS);
43    CSSR_PS_digital = digitalizacion(CSSR_PS);

45 % Representación gráfica de resultados (las métricas obtenidas
    % anteriormente) vs. los valores reales reportados por el nodo B. Se cargan
47 % las estadísticas con el script "carga_estadisticas" para poder tener en
    % el workspace los parámetros CSSR_CS y CSSR_PS y así poder representarlos
49 % gráficamente (recordar que estos parámetros se trabaja con ellos dentro
    % de la función "estudio_acc" pero cuando la función retorna se borran de
51 % memoria.

53    representaciones_graficas;

55 % Comparativa a nivel digital Acc. paquetes teórica vs. real

57    error_CS = 100*sum(xor(CSSR_CS_digital ,Acc_CS_digital))/length(Acc_CS_digital);
    disp('% Error en CS'); disp(error_CS);
59
    error_PS = 100*sum(xor(CSSR_PS_digital ,Acc_PS_digital))/length(Acc_PS_digital);
61    disp('% Error en PS'); disp(error_PS);

```

source/ejecutable.m

A.1.2. Estudio de la accesibilidad

```

function [Acc_CS_teorica ,Acc_PS_teorica ,Acc_CS_digital ,Acc_PS_digital ,M]=
    estudio_acc(NumEscenario)
2

4    % CALCULO DE ACC. EN EL DOWNLINK
    % =====
6
8    carga_estadisticas;

10   Link = 'DL';
    disp('Calculando Acc. Iub DL');
    [CSSR_CS_IUB_DL,CSSR_PS_IUB_DL]=ACCESS_IUB(NumEscenario ,Link ,Minutes_Voice3 ,
        Minutes_VT ,SHO_factor ,KB_PS_R99_DL,KB_PS_R99_UL,KB_HSDPA_DL_P6,KB_HSUPA,
        RAB_Att_PS_R99,RAB_Att_HSDPA,RAB_Att_HSUPA);
12
    M=[CSSR_CS_IUB_DL;CSSR_PS_IUB_DL]; %Matriz donde se van recogiendo las
        accesibilidades parciales
14
    %    % El tráfico que pasa a este subsistema no es el total , hay que restarle las
        llamadas bloqueadas
16 %
18    K_CS=CSSR_CS_IUB_DL; K_PS=CSSR_PS_IUB_DL; actualizar_trafico;

```

```

20      disp('Calculando Acc. CE DL');
      CSSR_CE_DL=ACCESS_CE(NumEscenario, Link, Minutes_Voice3, Minutes_VT, SHO_factor,
          KB_HSDPA_DL_P6, KB_PS_R99_DL, RAB_Att_HSDPA, RAB_Att_HSUPA, RAB_Att_PS_R99,
          KB_HSUPA, KB_PS_R99_UL);

22      M=[M;CSSR_CE_DL];

24      %      % El tráfico que pasa a este subsistema no es el total, hay que restarle las
          %      llamadas bloqueadas

26      K_CS=CSSR_CE_DL; K_PS=CSSR_CE_DL; actualizar_trafico;

28      disp('Calculando Acc. SF (DL)');
      %      SHO_factor=ones(1,96);
30      CSSR_SF_DL = ACCESS_SF(NumEscenario, Link, Minutes_Voice3, Minutes_VT, SHO_factor,
          KB_HSDPA_DL_P6, KB_PS_R99_DL, RAB_Att_HSDPA, RAB_Att_PS_R99);

32      M=[M;CSSR_SF_DL];

34      % Combinación de los factores

36      Acc_CS_DL = combina_factores(CSSR_CS_IUB_DL, CSSR_CE_DL, CSSR_SF_DL);
      Acc_PS_DL = combina_factores(CSSR_PS_IUB_DL, CSSR_CE_DL, CSSR_SF_DL);
38

40      % CALCULO DE ACC. EN EL UPLINK
42      % =====

44

46      Link = 'UL';

48      disp('Calculando Acc. CE UL');
      CSSR_CE_UL=ACCESS_CE(NumEscenario, Link, Minutes_Voice3, Minutes_VT, SHO_factor,
          KB_HSDPA_DL_P6, KB_PS_R99_DL, RAB_Att_HSDPA, RAB_Att_HSUPA, RAB_Att_PS_R99,
          KB_HSUPA, KB_PS_R99_UL);

50      M=[M;CSSR_CE_UL];

52      %      % El tráfico que pasa a este subsistema no es el total, hay que restarle
          %      las llamadas bloqueadas

54      K_CS=CSSR_CE_UL; K_PS=CSSR_CE_UL; actualizar_trafico;

56      disp('Calculando Acc. Iub UL');
      [CSSR_CS_IUB_UL, CSSR_PS_IUB_UL]=ACCESS_IUB(NumEscenario, Link, Minutes_Voice3,
          Minutes_VT, SHO_factor, KB_PS_R99_DL, KB_PS_R99_UL, KB_HSDPA_DL_P6, KB_HSUPA,
          RAB_Att_PS_R99, RAB_Att_HSDPA, RAB_Att_HSUPA);
58

60      M=[M;CSSR_CS_IUB_UL; CSSR_PS_IUB_UL];

62      % Combinación de los factores

64      Acc_CS_UL = combina_factores(CSSR_CE_UL, CSSR_CS_IUB_UL);
      Acc_PS_UL = combina_factores(CSSR_CE_UL, CSSR_PS_IUB_UL);

```

```

66      % COMPARACION CON LA ACC. REAL
68      % =====
70      % Valor comparable al reportado por el nodo
72      Acc_CS_teorica = 100*(Acc_CS_DL+Acc_CS_UL)/2;
74      Acc_PS_teorica = 100*(Acc_PS_DL+Acc_PS_UL)/2;
76      % DIGITALIZACIÓN DE LAS ACC. TEORICAS
78      % =====
80      % Cuando interesa evaluar únicamente la fiabilidad del modelo basándose
82      % en la calidad de la detección de degradación
84      umbral_digitalizado=70;
86      Acc_CS_digital = digitalizacion(Acc_CS_teorica);
88      Acc_PS_digital = digitalizacion(Acc_PS_teorica);
end

```

source/estudio_acc.m

A.2. Estudio de accesibilidades parciales

A.2.1. Modelo accesibilidad Iub

A.2.1.1. Nodos Dual-Stack

```

1  % Código ejecutable para la Tesis
3  % function
4  % [CSSR_CS_IUB,CSSR_PS_IUB]=ACCESS_IUB(NumEscenario , Link , Minutes_Voice , Minutes_VT ,
5  %   SHO_factor ,KB_PS_R99_DL,KB_PS_R99_UL,KB_HSDPA_DL_P6,KB_HSUPA, RAB_Att_PS_R99,
6  %   RAB_Att_HSDPA, RAB_Att_HSUPA)
7  function [CSSR_CS_IUB,CSSR_PS_IUB]=ACCESS_IUB(NumEscenario , Link , Minutes_Voice ,
8  %   Minutes_VT ,SHO_factor ,KB_PS_R99_DL,KB_PS_R99_UL,KB_HSDPA_DL_P6,KB_HSUPA,
9  %   RAB_Att_PS_R99,RAB_Att_HSDPA,RAB_Att_HSUPA)
11
12  carga_parametros;      % Carga de parámetros del Escenario
13  Factor = 'IUB';
14
15  % RECURSOS TOTALES EN EL SISTEMA (medido en bits por segundo)
16
17  C = 3 * 53 * 8;          % Medida en kbps (1 VCC)
18  C = C * NumVcc;         % Escalado para "n" VCCs
19
20  Recursos = modelado_recursos_CS(Link , Factor);      % Recursos Unitarios que
21  %   demanda cada flujo

```

```

19 [Trafico ,Erlangs] = conformado_trafico_CS (Factor ,Minutes_Voice ,Minutes_VT ,
    SHO_factor);
21 [m,n]=size ( Trafico );
    PB = zeros(m,n); BC = zeros(m,n);
    [PB,BC] = bloqueo_KR (Trafico ,Recursos ,C);
23
    PBloq = bloqueo_cuartos (Trafico ,BC); % Matriz ancho 96
25
    for k=1:1:size (PBloq ,1)
27         bloqueadas (k,:) = Erlangs (k,:) .* PBloq (k,:); % Matriz con n° de llamadas
            bloqueadas
    end
29
    BLOQUEO_CS_IUB = sum (bloqueadas) ./ sum (Erlangs);
31 BLOQUEO_CS_IUB ( isinf (BLOQUEO_CS_IUB)) = 0;
    BLOQUEO_CS_IUB ( isnan (BLOQUEO_CS_IUB)) = 0;
33
    CSSR_CS_IUB = 1 - BLOQUEO_CS_IUB;
35
37
39 %%% Estudio del bloqueo para el recurso de Iub para Paquetes
    %%% =====
41
43 % erlangs de tráfico
    switch (Link)
45     case 'DL'
        KB_HSPA = KB_HSDPA_DL_P6;
47         KB_PS_R99 = KB_PS_R99_DL;
            RAB_Att_HSPA = RAB_Att_HSDPA;
49         RAB_Att_PS384 = RAB_Att_PS_R99;
            C = CIP;
51         mu = 1e3;
            N = C/mu;
53     case 'UL'
        KB_HSPA = KB_HSUPA;
55         KB_PS_R99 = KB_PS_R99_UL;
            RAB_Att_HSPA = RAB_Att_HSUPA;
57         RAB_Att_PS384 = RAB_Att_PS_R99;
            CDL = C;
59         CE1 = 2.048e3; % Capacidad de un E1 en kbps
            CUL = NumE1 * CE1 - CDL;
61         mu = 100; % Kbps
            N = CUL/mu; % canales
63         N = round (N);
    end
65
    [T_PS,E_PS] = conformado_trafico_PS (Factor ,KB_HSPA,KB_PS_R99,RAB_Att_HSPA,
        RAB_Att_PS384);
67
    Trafico_PS = sum (T_PS); % Erlangs totales (HSPA + PS384)
69 Erlangs_PS = sum (E_PS);

```

```

71     for i=1:length(Trafico_PS)
72         PB_PS(i)= ErlangC(Trafico_PS(i),N);
73     end
74
75     PB_PS = bloqueo_cuartos(Trafico_PS,PB_PS);
77     CSSR_PS_IUB=1-PB_PS;
79 end

```

source/ACCESS_IUB.m

A.2.1.2. Modelo accesibilidad nodos Full-IP

Downlink

```

1  % Código ejecutable para la Tesis
3  function [CSSR_CS_IUB_DL,CSSR_PS_IUB_DL]=ACCESS_IUB_DL_FullIP(D,NumEscenario,media,
   varianza,CDL)
5  switch (D)
   case 'p',    x=1:1:15;
7               lambda=media;
               fdp=exp(-lambda)*lambda.^x./factorial(x);
9
   case 'n',    x=1:1:15;
11              fdp=1./sqrt(2*pi*varianza^2).*exp(-(x-media).^2./(2*varianza^2));
12  end
13
15  % Carga en memoria las estadísticas reales
   switch (NumEscenario)
17       case 7, cargar_datos_esc7;
18       case 8, cargar_datos_esc8;
19  end
21
22  % CALCULO DE LA ACCESIBILIDAD PARA CIRCUITOS
23  % =====
24  % En circuitos se tienen 3 tipos de tráfico que compiten por los recursos,
25  % que son la voz, la videollamada y la señalización.
26  %
27  %
28  % RECURSOS TOTALES EN EL SISTEMA (medido en bits por segundo)
29
30  CCS = 3 * 53 * 8; % Medida en kbps (1 VCC virtual);
31
32  % RECURSOS UNITARIOS QUE DEMANDA CADA FLUJO (medido en bits por segundo)
33
34  % b_Voice = 12.2e3;
35  % b_VT = 80e3;
   b_Voice = 12; % Medidos en Kbps y redondeo para evitar problemas en el algoritmo
   de Kaufman-Roberts
37  b_VT = 80;

```

```

39 Recursos = [b_Voice b_VT];
41 % ERLANGS DE TRÁFICO OFRECIDOS POR CADA FLUJO (Se incluyen los datos de la
% celda y se tiene en cuenta también el tráfico de soft-handover.
43
Erlangs_Voice = RAB_Att_CS.*SHO_factor;
45 Erlangs_VT = RAB_Att_VT.*SHO_factor;
47 % Reducción de la granularidad temporal, de 15 minutos a 1 minuto,
% suponiendo que esa cantidad se distribuye a lo largo de esos 15 minutos
49 % según una distribución de Poisson.
51 % Se construye una matriz, en cada fila la expansión de un cuarto de hora,
% y así hasta completar los 96 cuartos de hora que tiene un día.
53
MD_Voz = Erlangs_Voice'*fdp; % Matriz de distribución de tráfico de voz
55 MD_VT = Erlangs_VT'*fdp; % Matriz de distribución de tráfico de videollamada
57
% Se redimensiona todo en un solo vector, con el tráfico ahora minuto a
59 % minuto a lo largo de 24 horas (1440 minutos)
61 VD_Voz = reshape(MD_Voz',1,96*15); % Tráfico de voz minuto a minuto
VD_VT = reshape(MD_VT',1,96*15); % Tráfico de videollamada minuto a minuto
63
% Cálculo del bloqueo de cada flujo con el algoritmo de Kaufman-Roberts.
65 % Cada fila representa un minuto del día y cada columna un flujo distinto.
67 for i=1:1:1440
    Trafico=[VD_Voz(i) VD_VT(i)];
69     if (VD_Voz(i)<100/15)
        PB(i,:)=[0 0];
71     else
        PB(i,:)= kaufman_large(Trafico,Recursos,CCS);
73         if (isnan(sum(PB(i,:)))==1)
            disp('Error: un resultado es NaN');
75             disp('Tráfico'); disp(Trafico);
            disp('Recursos'); disp(Recursos);
77             disp('Capacidad'); disp(C);
        end
79     end
end
81
83 % Se representan los datos de accesibilidad sólo cada 15 minutos. Se
% calcula el bloqueo de cada cuarto de hora como la media del bloqueo de
85 % los 15 minutos anteriores.
87 BTVoz_minuto = PB(:,1)'; % Vector bloqueo teórico de voz cada minuto
BTVT_minuto = PB(:,2)';
89
for i=1:96
91     BTVoz_cuarto(i) = sum(BTVoz_minuto((i-1)*15+1:i*15))/15; % Bloqueo medio cada
15 min
    BTVT_cuarto(i) = sum(BTVT_minuto((i-1)*15+1:i*15))/15; % Bloqueo medio cada 15
min

```

```

93 end
95 %% ACCESIBILIDAD DE BLOQUEO POR Iub DE CIRCUITOS
97 CSSR_CS_IUB_DL =(1.-BTVoz_cuarto).*(1.-BTVT_cuarto);
99
101 %% Estudio del bloqueo para el recurso de Iub para Paquetes
103 %% =====
105 % En paquetes se tienen 2 tipo de tráfico , PS384 y HSDPA
107 % RECURSOS TOTALES EN EL SISTEMA (medido en kbits por segundo)
109 CPS=CDL - CCS;
111 mu = 1e3; % Kbps
112 N = CPS/mu; % canales
113 N = round (N); % Redondeo hasta el entero mas cercano para ser entero
114 % y no dar problemas en KR
115 % RECURSOS UNITARIOS QUE DEMANDA CADA FLUJO (medido en kbits por segundo)
117 b_HSDPA = KB_HSDPA_DL_P6.*(8/900); % 8 bits/byte y 900seg/15min
119 b_PS384DL = KB_PS_R99_DL.*(8/900);
120 b_TOTAL = b_HSDPA + b_PS384DL;
121 % erlangs de tráfico
123 Erlangs = b_TOTAL/mu;
125 % reducción de la granularidad temporal, de 15 minutos a 1 minuto,
126 % suponiendo que esa cantidad se distribuye a lo largo de esos 15 minutos
127 % según una distribución de poisson (fdp)
129 % se construye una matriz, en cada fila la expansión de un cuarto de hora,
130 % y así hasta completar los 96 cuartos de hora que tiene un día.
131 MD_IUBPS = Erlangs'*fdp; % matriz de distribución de tráfico de voz
133 % se redimensiona todo en un solo vector, con el tráfico ahora minuto a
134 % minuto a lo largo de 24 horas (1440 minutos)
135 VD_IUBPS = reshape(MD_IUBPS',1,96*15); % tráfico hsdpa minuto a minuto
137 % Cálculo del bloqueo de cada flujo con el algoritmo de Kaufman-Roberts.
138 % Cada fila representa un minuto del día y cada columna un flujo distinto.
139
140 for i=1:1:1440
141     PB(i,:) = ErlangC(VD_IUBPS(i),N);
142 end
143
144 % Se representan los datos de accesibilidad sólo cada 15 minutos. Se
145 % calcula el bloqueo de cada cuarto de hora como la media del bloqueo de
146 % los 15 minutos anteriores.
147
148 BTHSDPA_minuto = PB(:,1)'; % Vector bloqueo teórico de voz cada minuto

```



```

149 % Bloqueo medio cada 15 min
151 for i=1:96
      BTHSDPA_cuarto(i) = sum(BTHSDPA_minuto((i-1)*15+1:i*15))/15;
153 end
155 %%% ACCESIBILIDAD DE BLOQUEO POR Iub DE PAQUETES
CSSR_PS_IUB_DL = (1.-BTHSDPA_cuarto);
157
end

```

source/ACCESS_IUB_DL_FullIP.m

Uplink

```

2 function [CSSR_CS_IUB_UL,CSSR_PS_IUB_UL] = ACCESS_IUB_UL_FullIP(D, NumEscenario,
      media, varianza, CUL)
4 switch (D)
      case 'p', x=1:1:15;
6          lambda=media;
          fdp=exp(-lambda)*lambda.^x./factorial(x);
8
      case 'n', x=1:1:15;
10         fdp=1./sqrt(2*pi*varianza^2).*exp(-(x-media).^2./(2*varianza^2));
end
12
14 % Carga en memoria las estadísticas reales
% Carga en memoria las estadísticas reales
16 switch (NumEscenario)
      case 7, cargar_datos_esc7;
18         case 8, cargar_datos_esc8;
end
20
22 % CALCULO DE LA ACCESIBILIDAD IUB CS UL (ATM)
% =====
24 % En circuitos se tienen 3 tipos de tráfico que compiten por los recursos,
% que son la voz, la videollamada y la señalización.
26 %
%
28 % RECURSOS TOTALES EN EL SISTEMA (medido en bits por segundo)
30 % C = 3000 * 53 * 8 * 2;
CCS = 3 * 53 * 8; % Medida en kbps (1 VCC)
32
% RECURSOS UNITARIOS QUE DEMANDA CADA FLUJO (medido en bits por segundo)
34
% b_Voice = 12.2e3;
36 % b_VT = 64e3;
38 b_Voice = 12; % Medidos en Kbps y redondeo para evitar problemas en el algoritmo
      de Kaufman-Roberts
b_VT = 64;

```

```

40 Recursos = [b_Voice b_VT];
42
44 % ERLANGS DE TRÁFICO OFRECIDOS POR CADA FLUJO (Se incluyen los datos de la
44 % celda y se tiene en cuenta también el tráfico de soft-handover.
46 Erlangs_Voice = RAB_Att_CS.*SHO_factor;
46 Erlangs_VT = RAB_Att_VT.*SHO_factor;
48
48 % Reducción de la granularidad temporal, de 15 minutos a 1 minuto,
50 % suponiendo que esa cantidad se distribuye a lo largo de esos 15 minutos
50 % según una distribución de Poisson.
52 % Se construye una matriz, en cada fila la expansión de un cuarto de hora,
52 % y así hasta completar los 96 cuartos de hora que tiene un día.
54
54 MD_Voz = Erlangs_Voice'*fdp; % Matriz de distribución de tráfico de voz
56 MD_VT = Erlangs_VT'*fdp; % Matriz de distribución de tráfico de videollamada
58
58 % Se redimensiona todo en un solo vector, con el tráfico ahora minuto a
60 % minuto a lo largo de 24 horas (1440 minutos)
62
62 VD_Voz = reshape(MD_Voz',1,96*15); % Tráfico de voz minuto a minuto
62 VD_VT = reshape(MD_VT',1,96*15); % Tráfico de videollamada minuto a minuto
64
66 % Cálculo del bloqueo de cada flujo con el algoritmo de Kaufman-Roberts.
66 % Cada fila representa un minuto del día y cada columna un flujo distinto.
68
68 for i=1:1:1440
70     Trafico=[VD_Voz(i) VD_VT(i)];
70     if (VD_Voz(i)<100/15)
72         PB(i,:)= [0 0];
72     else
74         PB(i,:)= kaufman_large(Trafico,Recursos,CCS);
74         if (isnan(sum(PB(i,:)))==1)
76             disp('Error: un resultado es NaN');
76             disp('Tráfico'); disp(Trafico);
78             disp('Recursos'); disp(Recursos);
78             disp('Capacidad'); disp(CDL);
80         end
80     end
82 end
84
84 % Se representan los datos de accesibilidad sólo cada 15 minutos. Se
86 % calcula el bloqueo de cada cuarto de hora como la media del bloqueo de
86 % los 15 minutos anteriores.
88
88 BTVoz_minuto = PB(:,1)'; % Vector bloqueo teórico de voz cada minuto
90 BTVT_minuto = PB(:,2)';
92
92 for i=1:96
92     BTVoz_cuarto(i) = sum(BTVoz_minuto((i-1)*15+1:i*15))/15; % Bloqueo medio cada
94     15 min
94     BTVT_cuarto(i) = sum(BTVT_minuto((i-1)*15+1:i*15))/15;

```

```

end
96
%%% ACCESIBILIDAD DE BLOQUEO POR Iub DE CIRCUITOS
98 CSSR_CS_IUB_UL=(1.-BTVoz_cuarto).*(1.-BTVT_cuarto);
100
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
102 %%% Estudio del bloqueo para el recurso de Iub UL para Paquetes
104 %%% =====
106 % En paquetes se tienen 2 tipo de tráfico , PS384 y HSDPA
108 % RECURSOS TOTALES EN EL SISTEMA (medido en bits por segundo)
110 CPS = CUL - CCS;
112 mu = 1e3; % Kbps
114 N = CPS/mu; % canales
116 N = round(N); % Redondeo hasta el entero mas cercano para ser entero
118 % y no dar problemas en KR
120
122 % RECURSOS UNITARIOS QUE DEMANDA CADA FLUJO (medido en kbits por segundo)
124
126 b_HSUPA = KB_HSUPA.*(8/900); % 8 bits/byte y 900seg/15min
128 b_PS384UL = KB_PS_R99_UL.*(8/900);
130 b_TOTAL = b_HSUPA + b_PS384UL;
132
134 % erlangs de tráfico
136 Erlangs = b_TOTAL/mu;
138
140 % reducción de la granularidad temporal, de 15 minutos a 1 minuto,
142 % suponiendo que esa cantidad se distribuye a lo largo de esos 15 minutos
144 % según una distribución de poisson (fdp)
146
148 % se construye una matriz, en cada fila la expansión de un cuarto de hora,
150 % y así hasta completar los 96 cuartos de hora que tiene un día.
152 MD_IUBPS = Erlangs'*fdp; % matriz de distribución de tráfico de voz
154
156 % se redimensiona todo en un solo vector, con el tráfico ahora minuto a
158 % minuto a lo largo de 24 horas (1440 minutos)
160 VD_IUBPS = reshape(MD_IUBPS',1,96*15); % tráfico hsdpa minuto a minuto
162
164 % Cálculo del bloqueo de cada flujo con el algoritmo de Kaufman-Roberts.
166 % Cada fila representa un minuto del día y cada columna un flujo distinto.
168
170 for i=1:1:1440
172     PB(i,:) = ErlangC(VD_IUBPS(i),N);
174 end
176
178 % Se representan los datos de accesibilidad sólo cada 15 minutos. Se
180 % calcula el bloqueo de cada cuarto de hora como la media del bloqueo de
182 % los 15 minutos anteriores.
184
186 BTPSUL_minuto = PB(:,1)'; % Vector bloqueo teórico de voz cada minuto
188
190

```

```

% Bloqueo medio cada 15 min
152 for i=1:96
    BTPSUL_cuarto(i) = sum(BTPSUL_minuto((i-1)*15+1:i*15))/15;
154 end
156 %%% ACCESIBILIDAD DE BLOQUEO POR Iub DE PAQUETES
CSSR_PS_IUB_UL = (1.-BTPSUL_cuarto);
158
end

```

source/ACCESS_IUB_UL_FullIP.m

A.2.2. Modelo accesibilidad channel elements

```

1 function CSSR_CE = ACCESS_CE(NumEscenario, Link, Minutes_Voice, Minutes_VT, SHO_factor,
KB_HSDPA, KB_PS_R99_DL, RAB_Att_HSDPA, RAB_Att_HSUPA, RAB_Att_PS_R99, KB_HSUPA,
KB_PS_R99_UL)
3 carga_parametros; % Carga de parámetros del Escenario
5 Factor = 'CE';
7 % Calculo de la capacidad total del sistema (total de recursos a compartir)
9 switch(Link)
    case 'DL'
11     if (NumHSTX_60 ~= 0)
13         C_45 = 384; % Se parte de 384 CE para la tarjeta HSTX45
% C_45 = C_45-128; % Si se activa HSUPA: -128 CE
15         % C_45 = C_45-128; % Si se activa el 2º ResourceId: -128 CE
% C_45 = C_45-128; % Si se activa el 3º ResourceId: -128 CE
17
19         C_60 = 384; % Se parte de 384 CE para la tarjeta HSTX60
C_60 = C_60-0; % Si se activa HSUPA: -0 CE
C_60 = C_60-0; % Si se activa el 2º ResourceId: -0 CE
21         % C_60 = C_60-128; % Si se activa el 3º ResourceId: -128 CE
23
C_DL = NumHSTX_45*C_45 + NumHSTX_60*C_60; % Capacidad total del sistema
25
    else
27         C_45i = 384; % Se parte de 384 CE para la tarjeta HSTX45
C_45f = C_45i-128; % Si se activa HSUPA: -128 CE
C_45f = C_45f-128; % Si se activa el 2º ResourceId: -128 CE
29         C_DL = (NumHSTX_45-1)*C_45i + C_45f;
31     end
33     C_DL = C_DL*3; % Multiplica por 3, para funcionamiento correcto.
C = C_DL;
35     KB_HSPA = KB_HSDPA;
KB_PS_R99=KB_PS_R99_DL;
37     RAB_Att_HSPA = RAB_Att_HSDPA;
RAB_Att_PS384 = RAB_Att_PS_R99;
39 case 'UL'

```

```

41         C_13 = 128;           % Se parte de 128 CE para la tarjeta RAX13
42         C_14 = 128;           % Se parte de 384 CE para la tarjeta RAX14
43         C_UL = NumRAX13*C_13 + NumRAX14*C_14; % Capacidad total del sistema
44         C = C_UL;
45         KB_HSPA = KB_HSUPA;
46         KB_PS_R99=KB_PS_R99_UL;
47         RAB_Att_HSPA = RAB_Att_HSUPA;
48         RAB_Att_PS384 = RAB_Att_PS_R99;
49     end
50     Recursos = modelado_recursos_CS(Link, 'CE'); % Recursos Unitarios que
51         demanda cada flujo
52
53     [T_CS,E_CS] = conformado_trafico_CS( Factor , Minutes_Voice , Minutes_VT , SHO_factor)
54         ;
55     [T_PS,E_PS] = conformado_trafico_PS( Factor ,KB_HSPA,KB_PS_R99,RAB_Att_HSPA,
56         RAB_Att_PS384);
57
58     Trafico = [T_CS;T_PS];
59     Erlangs = [E_CS;E_PS];
60
61     [m,n]=size( Trafico );
62     PB = zeros(m,n); BC = zeros(m,n);
63
64     if NumEscenario==9
65         switch (Link)
66             case 'DL'
67                 C=65*3;
68             case 'UL'
69                 C=65*3;
70         end
71     end
72
73     Recursos = modelado_recursos_CS(Link, Factor); % Recursos Unitarios que
74         demanda cada flujo
75
76     [PB,BC] = bloqueo_KR( Trafico ,Recursos ,C);
77
78     PBloq = bloqueo_cuartos( Trafico ,BC); % Matriz ancho 96
79
80     for k=1:1:size(PBloq,1)
81         bloqueadas(k,:) = Erlangs(k,:) .* PBloq(k,:); % Matriz con n° de llamadas
82         bloqueadas
83     end
84
85     BLOQUEO_CE = sum(bloqueadas) ./ sum(Erlangs);
86     BLOQUEO_CE( isinf(BLOQUEO_CE) )=0;
87     BLOQUEO_CE( isnan(BLOQUEO_CE) )=0;
88
89     CSSR_CE=1-BLOQUEO_CE;
90 end

```

source/ACCESS_CE.m

A.2.3. Modelo accesibilidad códigos de canalización

```

1 function CSSR_SF = ACCESS_SF(NumEscenario,Link,Minutes_Voice,Minutes_VT,SHO_factor,
   KB_HSDPA,KB_PS_R99_DL,RAB_Att_HSPA,RAB_Att_PS_R99)
3
   carga_parametros;    % Carga de parámetros del Escenario (Se necesita la
                        % variable NumEscenario)
5
   Factor = 'SF';
7
   [T_CS,E_CS] = conformado_trafico_CS(Factor,Minutes_Voice,Minutes_VT,SHO_factor)
   ;
   [T_PS,E_PS] = conformado_trafico_PS(Factor,KB_HSDPA,KB_PS_R99_DL,RAB_Att_HSPA,
   RAB_Att_PS_R99);
9
   Trafico = [T_CS;T_PS];
11  Erlangs = [E_CS;E_PS];
13
   for i=1:length(Trafico)
15
       % Arbol inicial, n° codigos disponibles
       A1 = [1 2 4 8 16 32 64 128 256];
17       A2 = [1 2 4 8 16 32 64 128 256];
       A = A1 + A2;
19
       % Inicialmente tenemos 10 codigos de SF 16 ocupados para HSDPA
21       C_HSDPA = 20; % 2 arboles
       PB_HSDPA(i) = ErlangB(Trafico(4,i),C_HSDPA);
23       % comprobamos cuantos códigos quedan libres para los otro servicios
       v = [0 0 0 0 20 0 0 0 0];
25       Y = arbol(A,v);
27
       % Calculo probabilidad bloqueo en voz
       C_Voz = Y(8);
29       PB_voz(i) = ErlangB(Trafico(1,i),C_Voz);
31
       % Calculo de los codigos que quedan libres despues de utilizar el
       % servicio de voz
33       v = [0 0 0 0 0 0 0 0 0 ceil(Trafico(1,i))]; % Codigos utilizados voz
       Y = arbol(Y,v); % Devuelve n° codigos libres
35
       % Calculo probabilidad PS384
37       C_PS384 = Y(5); % Codigos de SF=32
       PB_PS384(i) = ErlangB(Trafico(3,i),C_PS384);
39
       PB_VT(i)=0;
41
   end
43
   PB=[PB_voz;PB_VT;PB_PS384;PB_HSDPA];
45
   BC = Trafico.*PB; % bloqueadas
47
   PBloq = bloqueo_cuartos(Trafico,BC);    % Matriz ancho 96
49
   bloqueadas = Erlangs.*PBloq;
51
   BLOQUEO_SF = sum(bloqueadas)./sum(Erlangs);

```

```

53 BLOQUEO_SF(isinf(BLOQUEO_SF))=0;
    BLOQUEO_SF(isnan(BLOQUEO_SF))=0;
55
    CSSR_SF=1-BLOQUEO_SF;
57
59 end

```

source/ACCESS_SF.m

A.2.4. Combinación de accesibilidades parciales

```

1 function P=combina_factores(A,B,C)
3     if nargin==3
        PA=1-A;
5         PB=1-B;
        PC=1-C;
7
        %
        P = 1-(PA+PB+PC-PA.*PB-PA.*PC-PB.*PC-2.*PA.*PB.*PC);
9         P = 1-(PA+PB+PC-PA.*PB-PA.*PC-PB.*PC+PA.*PB.*PC);
        end
11
        if nargin==2
13             PA=1-A;
                PB=1-B;
15
                P = 1-(PA+PB-PA.*PB);
17
        end
19 end

```

source/combina_factores.m

A.2.5. Digitalización de accesibilidad teórica

```

function y = digitalizacion(x)
2
    y=zeros(1,length(x));
4
    m = 0.99*mean(x);
6
    y = x>m;
8
end

```

source/digitalizacion.m

A.3. Estudio de bloqueo en redes multiservicio

A.3.1. Modelado de recursos CS

```

1  function Recursos = modelado_recursos_CS(Link,Factor)
3
4  if strcmp( Factor , 'IUB')==1
5      b_Voice = 12;    % Medidos en Kbps y redondeo para evitar problemas en el
                        % algoritmo de Kaufman-Roberts
6      b_VT = 64;
7      b_Signal = 6;
8      Recursos = [b_Voice b_VT b_Signal];
9  end
11
12 if strcmp( Factor , 'CE')==1
13     switch (Link)
14         case 'DL'
15             b_Voice_DL = 1;    % Demanda unitaria de CE del servicio de Voz
16             b_VT_DL = 4;      % Demanda unitaria de CE del servicio
                                % Videollamada (CS64)
17             b_PS384_DL = 16;  % Demanda unitaria de CE para conmutación de
                                % paquetes (PS384)
18             b_HSDPA = 1.3;    % Demanda unitaria de CE para HSDPA (21 CE cada
                                % 16 usuarios)
19
20             % Dado que b_HSDPA es un numero decimal, multiplicamos por 3 todo (
                                % trafico ,
                                % demanda, capacidad... ) para que funcione el algoritmo de Kaufman
                                % -Roberts
21
22             b_Voice_DL = b_Voice_DL*3;
23             b_VT_DL = b_VT_DL*3;
24             b_PS384_DL = b_PS384_DL*3;
25             b_HSDPA = 4;
26             Recursos_DL = [b_Voice_DL b_VT_DL b_PS384_DL b_HSDPA];
27             Recursos = Recursos_DL;
28
29         case 'UL'
30             b_Voice_UL = 1;    % Demanda unitaria de CE del servicio de Voz
31             b_VT_UL = 2;      % Demanda unitaria de CE del servicio
                                % Videollamada (CS64)
32             b_PS384_UL = 8;   % Demanda unitaria de CE para conmutación de
                                % paquetes (PS384)
33             b_HSUPA = 2;     % Demanda unitaria de CE para HSUPA (32 CE cada
                                % 16 usuarios)
34
35             % Dado que b_HSUPA es un numero decimal, multiplicamos por 3 todo (
                                % trafico ,
                                % demanda, capacidad... ) para que funcione el algoritmo de Kaufman
                                % -Roberts
36
37             Recursos_UL = [b_Voice_UL b_VT_UL b_PS384_UL b_HSUPA];
38             Recursos = Recursos_UL;
39     end
40 end
41
42 if strcmp( Factor , 'SF')==1
43     % Los codigos de Spreading se definen en forma de arbol, en cada nivel hay
44     % diferente Factor de Spreading (SF).

```



```

47     % Se produce una demanda unitaria de SF dependiendo del tipo de servicio
48     disp('Se están analizando recursos para SF')
49     b_Voice_DL = 256;    % SF del servicio de Voz
50     b_VT_DL = 32;      % SF del servicio Videollamada (CS64)
51     b_PS384_DL = 32;   % SF para conmutación de paquetes (PS384)
52     b_HSDPA = 16;      % SF para HSDPA (21 CE cada 16 usuarios)
53     Recursos = [b_Voice_DL b_VT_DL b_PS384_DL b_HSDPA];
54 end
55 end

```

source/modelado_recursos_CS.m

A.3.2. Modelado de recursos PS

```

1  function Recursos = modelado_recursos_PS(Link)
3
4  %     b_Voice = 12;    % Medidos en Kbps y redondeo para evitar problemas en el
5  %     algoritmo de Kaufman-Roberts
6  %     b_VT = 64;
7  %
8  %     switch(Link)
9  %         case 'DL'
10 %             b_Signal = 6;
11 %             Recursos = [b_Voice b_VT b_Signal];
12 %         case 'UL'
13 %             Recursos = [b_Voice b_VT];
14 %     end
15 end

```

source/modelado_recursos_PS.m

A.3.3. Conformado de tráfico CS

```

1  % Función de conformado de tráfico
2  %
3  % Sintaxis:
4  %     Trafico = conformado_trafico_CS(Minutes_Voice, Minutes_VT, SHO_factor)
5  %
6  % Devuelve:
7  %     Trafico, matriz con tráficos interpolados y Erlangs, matriz con tráficos sin
8  %     interpolar.
9  function [Trafico, Erlangs] = conformado_trafico_CS(Factor, Minutes_Voice, Minutes_VT,
10     SHO_factor)
11
12     % ERLANGS DE TRÁFICO OFRECIDOS POR CADA FLUJO (Se incluyen los datos de la
13     % celda y se tiene en cuenta también el tráfico de soft-handover.
14
15     Erlangs_Voice = Minutes_Voice.*SHO_factor/15;
16     Erlangs_VT = Minutes_VT.*SHO_factor/15;
17     Erlangs_Signal = Erlangs_Voice + Erlangs_VT;

```

```

17   VD_Voz = interpolacion_trafico(Erlangs_Voice,1,15);
19   VD_VT = interpolacion_trafico(Erlangs_VT,1,15);
21   VD_Sign = interpolacion_trafico(Erlangs_Signal,1,15);
23   if strcmp( Factor , 'IUB' )
24       Trafico = [VD_Voz;VD_VT;VD_Sign];
25       Erlangs = [Erlangs_Voice;Erlangs_VT;Erlangs_Signal];
26   else
27       Trafico = [VD_Voz;VD_VT];
28       Erlangs = [Erlangs_Voice;Erlangs_VT];
29   end
end

```

source/conformado_trafico_CS.m

A.3.4. Conformado de tráfico PS

```

% Función de conformado de tráfico
2 %
% Sintaxis:
4 %   Trafico = conformado_trafico_PS( Factor ,KB_HSPA,KB_PS_R99,RAB_Att_HSPA,
   RAB_Att_PS384)
6 function [Trafico,Erlangs] = conformado_trafico_PS( Factor ,KB_HSPA,KB_PS_R99,
   RAB_Att_HSPA,RAB_Att_PS384)
8   if strcmp( Factor , 'IUB' )
9       lambda_HSPA = KB_HSPA.*(8/900);    % 8 bits/byte y 900seg/15min
10      lambda_PS384 = KB_PS_R99.*(8/900); % (paso a Kbps)
11      mu=1e3; % Kbps
12
13      Erlangs_HSPA = lambda_HSPA/mu;
14      Erlangs_PS384 = lambda_PS384/mu;
15   else
16       Erlangs_HSPA = RAB_Att_HSPA;
17       Erlangs_PS384 = RAB_Att_PS384;
18   end
20   VD_HSPA = interpolacion_trafico(Erlangs_HSPA,1,15);
21   VD_PS384 = interpolacion_trafico(Erlangs_PS384,1,15);
22
23   Erlangs = [Erlangs_PS384;Erlangs_HSPA];
24   Trafico = [VD_PS384;VD_HSPA];
26 end

```

source/conformado_trafico_PS.m

A.3.5. Algoritmo original K-R

```

% Esta función implementa el algoritmo de J.S.Kaufman para el cálculo de
2 % bloqueo en redes, donde se comparten "C" unidades de un recurso entre "k"

```

```

% clientes , los cuales demandan cada uno diferentes unidades del mismo
4 % recurso.
%
6 % Sintaxis: PB=kaufman(a,b,c)
% - PB(i) es la probabilidad de bloqueo para los clientes de clase "i".
8 % - a(i) es el tráfico ofrecido de clientes de clase "i".
% - b(i) es la cantidad de recursos que consume cada cliene de clase "i".
10 % - C es el no total de unidades que se tienen del recurso compartido.

12 function PB = kaufman(a,b,C)

14     k=length(a);          % "k" tipos de clientes
    V=[1 zeros(1,C)];      % Es la función q^ inicializada a 0s y q^(0)=1

16
18     for j=1:1:C
19         B=0;
20         for i=1:1:k          %Comprobamos si el argumento de q^<0, que vale 0
21             if(j-b(i)<0)
22                 Z=0;
23             else
24                 Z=V(j-b(i)+1);
25             end
26             B=B+a(i)*b(i)*Z;
27         end
28         V(j+1)=(1/j)*B;
29     end

30     q0=1/sum(V);
31
32     q=q0.*V;
33
34     % Cálculo de las probabilidades de bloqueo para cada tipo de cliente "i"
35
36     PB=zeros(1,k);
37
38     for i=1:1:k
39         n2=b(i)-1;
40         PB(i)=sum(q(end-n2:end));
41     end
42
end

```

source/kaufman.m

A.3.6. Función Distribución K-R para el algoritmo original

```

function [q] = krDistribution(a,b,C)
2
k=length(a);
4 q=zeros(1,C);

6 for j=1:C
7     for i=1:k
8         if(j-b(i)>=0)
9             if(j-b(i)==0)

```

```

10         q(j) = q(j)+a(i)*b(i);
           else
12         q(j) = q(j)+a(i)*b(i)*q(j-b(i));
           end
14     end
       end
16
       if(q(j)>0)
18         q(j) = (1/j)*q(j);
       end
20 end
22 q0 = 1/(sum([1 q]));
   q = [q0 q0.*q];

```

source/krDistribution.m

A.3.7. Función algoritmo K-R empleando FFT para evitar desbordamiento

```

1 % Esta función implementa una mejora del algoritmo de Kaufman–Roberts para
2 % problemas de números muy grandes, en los que tradicionalmente la
3 % implementación original se desborda.
5 function PB=kaufman_large(a,b,C)
7     k=length(b); %Nº de clases de tráfico
9     % Paso 1) Calcular e_max utilizando la expresión (9) del paper.
   E=realmax;
11 %     emax=(1/C)*(E/C^2)^(1/k);
   emax=1; %Para hacer pruebas, el resultado es independiente de emax
13
   % Paso 2) Calcular la función de distribución de cada clase de tráfico (4) y
   (5)
15 %     =====
17 nkmax=floor(C./b); % Máximo de usuarios que puede haber de tipo "k".
19 p=zeros(k,C+1); %Cada fila de esta matriz, una función de distribución
21 for j=1:k
   if (a(j)<nkmax(j)) % Función de distribución es convexa
23     na=floor(a(j)); % Máximo de la función convexa
   p(j,na*b(j)+1)=emax; % Pongo el valor en ese punto
25     for s=(na+1)*b(j):b(j):C+1 % Calculo los valores a la derecha del pto
   máximo, de izquierda a derecha
       nk=s/b(j);
27     p(j,s+1)=p(j,(nk-1)*b(j)+1)*a(j)/(nk+1);
   end
29
   for s=(na-1)*b(j):-b(j):0 % Calculo los valores a la izquierda del
   pto. máximo, de derecha a izquierda
31     nk=s/b(j);
   p(j,s+1)=p(j,(nk+1)*b(j)+1)*(nk+1)/a(j);
33 end

```

```

35     else                                     % Función de distribución es creciente
        p(j,nkmax(j)*b(j)+1)=emax; % El máximo a la derecha
        for s=C-b(j):-b(j):0 % Calculo los valores hacia la izquierda, de
            derecha a izquierda
37             nk=s/b(j);
            p(j,s+1)=p(j,(nk+1)*b(j)+1)*(nk+1)/a(j);
39         end
        end
41     end

43
45     % Paso 3) Se calcula la convolución de todas las distribuciones (7) con
        % las transformadas de Fourier

47     Z=p(1,:);
        for i=2:1:k
49         fun1=Z;
            fun2=p(i,:);
51 %         Z=ifft(fft(fun1).*fft(fun2)); %Esta línea y la siguiente hacen lo mismo
            Z=conv(fun1,fun2);
53         Z=Z(1:C+1);

55     end

57     G=Z;

59
61     % Paso 4) Cálculo de las probabilidades de bloqueo, con ecuación (1)

        D=sum(G); % Denominador, es común
63
        PB=ones(1,k); % Inicializamos el vector
65
        for s=1:1:k
67             Nd=G(C-b(s)+2:C+1);
                N=sum(Nd);
69             PB(s)=N/D;
        end
71 end

```

source/kaufman_large.m

A.3.8. Función Erlang-B

```

1 %%% Formula de Erlang B
  % A: cantidad total de trafico en Erlangs
3 % m: numero de recursos
  % PB: probabilidad de bloqueo
5
  % function PB=ErlangB(A,m)
7 %     if (A == 0)
  %         PB = 0;
9 %     else
  %         InvB = 1;
11 %         for k=1:m

```

```

%           InvB = 1 + k/(A*InvB);
13 %       end
%       PB = 1/InvB;
15 %   end
% end
17
19
function PB=ErlangB(A,m)
21     B = A/(1+A);
     for k=2:1:m
23         B = 1/(1 + k/(A*B));
     end
25     PB = B;
end

```

source/ErlangB.m

A.3.9. Función Erlanc-C

```

function [ PB ] = ErlangC( A, N )
2 %ErlangC
% A: Intensidad total de trafico en Erlangs.
4 % N: numero de servidores
% output: PW Probabilidad de bloqueo.
6
     numerador = (A^N)*N/(factorial(N)*(N-A));
8     b=0;
     for i=0:1:N-1
10         b = (A^i)/factorial(i) + b;
     end
12     denominador = b + numerador;
     PB = numerador/denominador;
14
%     PB = ErlangB(A,N);
16 %     PW = N*PB/(N-A*(1-PB));
18 end

```

source/ErlangC.m

A.4. Otras funciones

A.4.1. Función constructor de árbol

```

1 % Función que analiza qué sucede con un determinado árbol de códigos "A"
% cuando se utiliza un código de longitud "n".
3 %
% Argumentos:
5 %     A: vector "árbol", contiene los códigos disponibles de cada SF
%         Cada posición de A corresponde a un SF.
7 %     v: vector con los códigos que se utilizarán de cada SF.
%

```

```

9 % Ejemplo de árbol de códigos:
%
11 % A=[1 2 4 8 16] significa que tenemos:
%
13 % 1 código de SF=1
% 2 códigos de SF=2
15 % 4 códigos de SF=4
% 8 códigos de SF=8
17 % 16 códigos de SF=16.
%
19 % (este es un ejemplo de árbol "virgen", sin códigos utilizados).
%
21 % La función devuelve un vector "Y", árbol de salida con los códigos
% disponibles para una futura iteración.
23
function y=arbol(A,v)
25
L=length(A); % Nº de niveles del árbol
27
% Primero, dejan de estar disponibles los códigos que se utilizan
29
y=A-v;
31
% for k=1:L
33 % if (y(k)<0)
% disp('codigos insuficientes');
35 % end
% end
37
39
% Bloqueo de códigos ascendientes
41
for k=1:L
43 if (v(k)>0)
r=[(1+floor(v(k)/2))*ones(1,k-1) zeros(1,L-k+1)];
45 y=y-r;
end
47 end
49
% Bloqueo de códigos descendientes
51
for k=1:L-1
53 if (v(k)>0)
for t=1:L-k
55 y(k+t)=y(k+t)-(2*t)*v(k);
end
57 end
59 end
end

```

A.4.2. Actualización de tráfico entrantes a cada subsistema

A.4.2.1. Tráficos CS

```

2  % Función que, elemento a elemento, multiplica v(i) por k(i)
4  RAB_Att_CS      = RAB_Att_CS.*k;
   RAB_Att_VT      = RAB_Att_VT.*k;

```

source/actualizar_datos_CS.m

A.4.2.2. Tráficos PS

```

2  % Función que, elemento a elemento, multiplica v(i) por k(i)
4
6  RAB_Att_HSDPA   = RAB_Att_HSDPA.*k ;
   RAB_Att_PS_R99  = RAB_Att_PS_R99.*k ;
   RAB_Att_HSUPA   = RAB_Att_HSUPA.*k ;

```

source/actualizar_datos_PS.m

A.4.3. Reconstrucción de tráfico a intervalos de 15 minutos

```

1  function B=bloqueo_cuartos(Trafico,BC)
3
5     [m,n]=size(Trafico);
7
9     if n==1
11        B=Trafico./BC;
13
15    else
17        if n<15
19            grupo = n;
21        else
23            grupo=15;
25        end
27
29        B=zeros(m,n/grupo);
31
33        BCrs=reshape(BC,m,grupo,n/grupo);
35        TRAFrs=reshape(Trafico,m,grupo,n/grupo);
37        B=reshape(sum(BCrs,2)./sum(TRAFrs,2),m,n/grupo);
39    end
41
43    B(isinf(B))=0;
45    B(isnan(B))=0;
47
49 end

```

source/bloqueo_cuartos.m

A.4.4. Carga de parámetros de los escenarios

```

1  % Este fichero carga en memoria los parámetros de entrada para el modelo
3  switch (NumEscenario)
4      case 0
5          D='p';
6          NumEscenario=0;
7          NumVcc = 1;
8          CIP = 23e3; %Kbps
9          NumE1 = 2;
10         NumHSTX_45=1;
11         NumHSTX_60=1;
12         NumRAX13 = 2;
13         NumRAX14 = 0;
14         umbral_discretizado=0.99;
15         m=ceil(0.02*1440);
16     case 1
17         D='p';
18         NumEscenario=1;
19         NumVcc = 1;
20         CIP = 23e3; %Kbps
21         NumE1 = 5;
22         NumHSTX_45=1;
23         NumHSTX_60=1;
24         NumRAX13 = 2;
25         NumRAX14 = 0;
26         umbral_discretizado=0.99;
27         m=ceil(0.02*1440);
28     case 2,
29         disp('Escenario 2');
30         D='p';
31         NumEscenario=2;
32         NumVcc = 2;
33         CIP = 23e3; %Kbps
34         NumE1 = 5;
35         NumHSTX_45=1;
36         NumHSTX_60=1;
37         NumRAX13 = 2;
38         NumRAX14 = 0;
39         umbral_discretizado=0.99;
40         m=ceil(0.02*1440);
41     case 3,
42         D='p';
43         NumEscenario=3;
44         NumVcc = 1;
45         CIP = 23e3; %Kbps
46         NumE1 = 5;
47         NumHSTX_45=1;
48         NumHSTX_60=1;
49         NumRAX13 = 2;
50         NumRAX14 = 0;
51         umbral_discretizado=0.99;
52         m=ceil(0.02*1440);
53     case 4,

```

```

55     D='p';
        NumEscenario=4;
57     NumVcc = 2;
        CIP = 23e3; %Kbps
59     NumE1 = 5;
        NumHSTX_45=1;
61     NumHSTX_60=1;
        NumRAX13 = 2;
63     NumRAX14 = 0;
        umbral_discretizado=0.99;
65     m=ceil(0.02*1440);
    case 5,
67     D='p';
        NumEscenario=5;
69     NumVcc = 1;
        CIP = 30e3; %Kbps
71     NumE1 = 1;
        NumHSTX_45=2;
73     NumHSTX_60=0;
        NumRAX13 = 0;
75     NumRAX14 = 2;
        umbral_discretizado=0.99;
77     m=ceil(0.02*1440);
    case 6,
79     D='p';
        NumEscenario=6;
81     NumVcc = 1;
        CIP = 30e3; %Kbps
83     NumE1 = 1;
        NumHSTX_45=2;
85     NumHSTX_60=0;
        NumRAX13 = 0;
87     NumRAX14 = 3;
        umbral_discretizado=0.99;
89     m=ceil(0.02*1440);
    case 9,
91     D='p';
        NumEscenario=9;
93     NumVcc = 1;
        CIP = 30e3; %Kbps
95     NumE1 = 2;
        NumHSTX_45=0; % Para conservar la equivalencia , ya que las tarjetas del
        Nodo son más antiguas
97     NumHSTX_60=0;
        NumRAX13 = 0;
99     NumRAX14 = 0;
        umbral_discretizado=0.99;
101    m=ceil(0.02*1440);
    case 10,
103    D='p';
        NumEscenario=10;
105    NumVcc = 1;
        CIP = 30e3; %Kbps
107    NumE1 = 2;
        NumHSTX_45=1;
109    NumHSTX_60=0;

```

```

111     NumRAX13 = 2;
112     NumRAX14 = 0;
113     umbral_discretizado=0.99;
114     m=ceil(0.02*1440);
115     case 11,
116         D='p';
117         NumEscenario=1;
118         NumVcc = 1;
119         CIP = 23e3; %Kbps
120         NumE1 = 5;
121         NumHSTX_45=1;
122         NumHSTX_60=1;
123         NumRAX13 = 2;
124         NumRAX14 = 0;
125         umbral_discretizado=0.99;
126         m=ceil(0.02*1440);
127     end

```

source/carga_parametros.m

A.4.5. Calibrado de tráfico (parámetro λ)

```

% Implementación de un sistema realimentado para la calibración del
2 % tráfico
% Lo es el valor "actual" de lambda, que viene dado por cálculos en
4 % escenarios anteriores.

6 function L = calibrado_trafico(Lo,D,NumEscenario,numVcc,CIP,NumE1,NumHSTX_45,
    NumHSTX_60,NumRAX13,NumRAX14,umbral_discretizado,m)

8     f=0;
9     k=Lo;
10    carga_estadisticas;
11    t=0; % Contador auxiliar del nº de iteraciones que llevamos
12    aux = 0; % Variable auxiliar para saber si ya se encontró una lambda de
    calidad

14    while f==0
15        disp('Iteración para lambda'),disp(k);
16        if k>=5 % Se establece cota superior para lambda (fidelidad del
    enventanado de tráfico)
17            sprintf('El valor de lambda = %g sobrepasa la cota superior prefijada',
    k)
18            L=(k+Lo)/2;
19            f=1;
20        else
21            lambda=k;
22            media = lambda;
23            varianza = lambda;
24            CSSR_T = ACC_Teorica(D,NumEscenario,media,varianza,numVcc,CIP,NumE1,
    NumHSTX_45,NumHSTX_60,NumRAX13,NumRAX14);
25            Acc_td = discretizado(CSSR_T,0.7);
26            CSSR_CSd = discretizado(CSSR_CS,umbral_discretizado);
    DELTA = xor(Acc_td,CSSR_CSd); % Compara accesibilidad real con
    teórica.

```

```

28     FC = sum(DELTA)                                % Número de discrepancias en el
           calibrado de tráfico (cuento el nº de 1s en el vector)

30     m=10;
     f = comprueba_calidad(FC,m);    % Comprueba la calidad del modelo así
           implementado

32     if f==0 && aux==0    % Si el sistema no es de calidad, incremento
           lambda en pasos de 0.5
34         sprintf('lambda = %g no proporciona un sistema de calidad. Se va a
           probar con lambda = %g ',k,k+0.5)
           k=k+0.5;
36     end

38     if f==1            % Si el sistema es de calidad, busco si una lambda más
           pequeña también da un sistema con calidad
           sprintf('lambda = %g proporciona un sistema de calidad. Se va a
           probar con lambda = %g ',k,k-0.5)
40         lambda_aux=k;
           L=(k+Lo)/2; % Se aprovechan las experiencias anteriores (Lo).
42         k=k-0.5;
           aux = 1;    % "Flag" que indica que ya se encontró una lambda de
           calidad (para no repetir el cálculo si después f==0
44         f=0;        % Vuelvo a poner f a 0 para que vuelva a entrar en el
           bucle while

           end

46     if f==0 && aux==1    % Si la lambda no es de calidad pero ya había
           encontrado otra, termino
48         sprintf('lambda = %g no proporciona un sistema de calidad, pero
           lambda = %g sí.',k,lambda_aux)
           sprintf('En base a experiencias anteriores, se determina que lambda
           = %g ',L)
50         f=1;

           end
52     end
54 end

```

source/calibrado_trafico.m

A.4.6. Representaciones gráficas

```

% Comparativa a nivel analógico Acc. circuitos teórica vs. real
2 % close all
   figure
4   plot(CSSR_CS, 'k-');
   hold on; grid on;
6   plot(Acc_CS_teorica, 'r');
   title('Comparativa Accesibilidad CS');
8   legend('Acc. real', 'Acc. teórica', 'Location', 'SouthWest');
   xlabel('Tiempo (1/4 hora)');
10  ylabel('% Accesibilidad');

12 % Comparativa a nivel analógico Acc. paquetes teórica vs. real

```

```
14 figure
15 plot(CSSR_PS, 'k-');
16 hold on; grid on;
17 plot(Acc_PS_teorica, 'r');
18 title('Comparativa Accesibilidad PS');
19 legend('Acc. real', 'Acc. teórica', 'Location', 'SouthWest');
20 xlabel('Tiempo (1/4 hora)');
21 ylabel('% Accesibilidad');
22
23 % Comparativas a nivel digital
24
25 % Comparativa a nivel digital Acc. circuitos teórica vs. real
26 figure
27 stem(CSSR_CS_digital);
28 hold on;
29 stem(-Acc_CS_digital, 'r');
30 title('Digitalización de señales Acc CS teórica vs. real');
31 legend('Digital Real', 'Digital Teórica');
32
33
34
35 % Comparativa a nivel digital Acc. paquetes teórica vs. real
36 figure
37 stem(CSSR_PS_digital);
38 hold on;
39 stem(-Acc_PS_digital, 'r');
40 title('Digitalización de señales Acc PS teórica vs. real');
41 legend('Digital Real', 'Digital Teórica');
```

source/representaciones_graficas.m

Apéndice B

ATM (*Asynchronous Transfer Mode*)

B.1. Introducción

ATM (*Asynchronous Transfer Mode*) es una tecnología de transporte ampliamente empleada en redes de telecomunicaciones y que opera a nivel de capa 2 según el modelo de referencia OSI.

ATM es una tecnología orientada a conexión, diseñada para el transporte de múltiples tipos de servicios entre ellos voz, vídeo y datos a través de redes públicas y privadas y provee gran escalabilidad del ancho de banda.

B.2. Conceptos

B.2.1. Celdas

ATM se basa en la multiplexación de celdas. Son pequeños paquetes de tamaño fijo, estandarizadas a 53 bytes, figura B.1. De estos 53 bytes los 5 primeros se corresponden con datos de cabecera y el resto con la carga de datos a transportar. El hecho de que las celdas sean de tamaño fijo facilita el proceso de conmutación al no tener que verificar el tamaño variable de celdas. No hay que incluir en el encabezado el tamaño de la celda. El porqué de los 53 bytes se explica debido a que celdas más grandes provocan tiempos de retardo en la transmisión que servicios como voz y vídeo no pueden tolerar, mientras que celdas más pequeñas reducen la eficiencia en el manejo del ancho de banda.

ATM toma los paquetes o tramas de los servicios a transportar y los divide en pequeños segmentos que multiplexa a través de una línea.

El encargado de segmentar los datos es la capa de adaptación ATM o AAL.

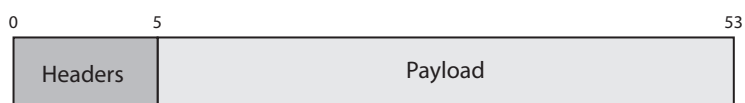


Figura B.1: Celda ATM

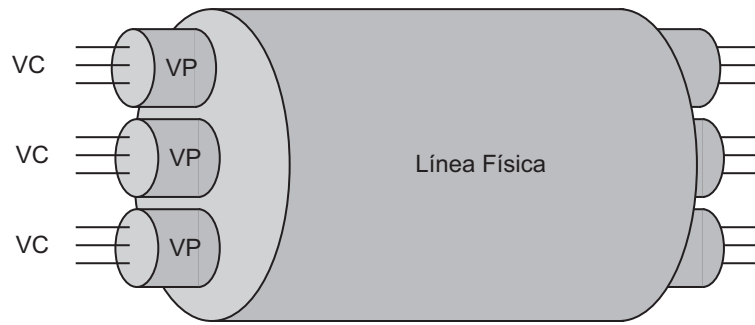


Figura B.2: Conexiones virtuales en ATM

B.2.2. Conexiones Virtuales

Las conexiones virtuales son conexiones lógicas para establecer comunicación entre dos clientes de la red ATM. Varias conexiones virtuales subdividen una sola interfaz física.

El concepto de canal virtual representa una conexión entre dos puntos de la red ATM mientras que la trayectoria virtual o *virtual path* representa un conjunto de canales virtuales (ver figura B.2).

B.2.3. Conmutación en ATM

A la hora de hablar de tipos de conmutación pueden encontrarse dos tipos:

B.2.3.1. Conmutación de circuitos

Dedica un circuito físico a cada comunicación. Su mayor ventaja es poseer tiempos de retraso pequeños y estables en la transmisión, además de ofrecer un ancho de banda garantizado a los usuarios. En contrapartida las conexiones son poco flexibles y desaprovecha gran parte del ancho de banda de la línea utilizada.

B.2.3.2. Conmutación de paquetes

Se comparte el medio físico, tal y como sucede por ejemplo en la tecnología IP. Su mayor ventaja es que aprovecha al máximo el ancho de banda y dota de mayor flexibilidad a las conexiones. En contrapartida no se garantiza el ancho de banda y el retraso de los paquetes es variable.

ATM combina las ventajas de ambos métodos de conmutación:

- ⇒ Ancho de banda garantizado y el tiempo de retraso estable de la conmutación de circuitos.
- ⇒ Aprovechamiento del ancho de banda y la flexibilidad en las conexiones.

B.3. Calidad de Servicio

ATM implementa un fuerte control sobre las conexiones a fin de dotar de calidad de servicio a cada una. Para ello, cuando se establece una nueva conexión se inicia un «contrato» entre la red y el usuario por el que ambos se comprometen a un intercambio de información dado. La red reserva los recursos necesarios y el usuario se compromete a transmitir de acuerdo a unos criterios:

- ⇒ Se especifican parámetros para el flujo de datos de una conexión.

⇒ Establece un compromiso de la red hacia las terminales.

La calidad de servicio vendrá dada por el tipo de tráfico que se curse. Existen diferentes categorías:

⇒ Flujo constante CBR:

- Ancho de banda estable y tiempo de retraso garantizado.
- Transporta servicios que requieren señal de sincronía.
- Utilizado para voz y vídeo digitales y emulación de circuitos.

⇒ Flujo variable - Tiempo Real VBR-RT:

- Provee ancho de banda variable por demanda a servicios susceptibles al tiempo de transmisión.
- Utilizado para voz y vídeo «paquetizados».

⇒ Flujo variable VBR-NRT:

- Provee ancho de banda variable por demanda a servicios menos susceptibles al tiempo de transmisión.

⇒ Inespecificado UBR:

- Provee el mejor trato posible sin garantías de ninguna especie.

⇒ Flujo disponible ABR:

- Provee de un ancho de banda variable de acuerdo a la congestión de la red. \item Se utiliza para transportar servicios muy flexibles al tiempo de transmisión y al ancho de banda como IP y IPX.
- Requiere de Control de Flujo a lazo cerrado.

B.4. Capas de Adaptación

AAL nace como un mecanismo de la tecnología ATM para la adaptación del formato de celda desde la capa de enlace de datos a los niveles superiores de otras arquitecturas, tales como TCP/IP o el modelo OSI. Está definido en el ITU y se diseñaron 5 protocolos adaptados a las diferentes arquitecturas de red que pueden soportar ATM.

Constan de dos partes:

⇒ Subcapa de convergencia o CS (*Convergence Sub-layer*).

⇒ Subcapa de segmentación y reensamble o SAR (*Segmentation and Reassembly Sub-layer*).

B.4.1. Capa AAL1

Se emplea habitualmente con conexiones que transporten un flujo constante de bits, por lo que es idónea para el transporte de voz y vídeo. Transporta servicios que requiera sincronización entre emisor y receptor y además puede emular circuitos.

B.4.2. Capa AAL2

AAL Tipo 2 soporta rt-VBR (*Variable Bit Rate in real time*), de circuito orientado a la conexión de tráfico síncrono. Su servicio es de baja prioridad y garantizado. Se utiliza en compresión de vídeo.

B.4.3. Capa AAL3-4

Soportan al VBR, tráfico de datos, circuitos orientados a la conexión, tráfico asíncrono (por ejemplo X.25 de datos) o a los paquetes de datos no orientados a la conexión (ej:tráfico SMDS) con una cabecera (*header*) adicional de 4 bytes en el *payload* de la celda. Por ejemplo, Frame Relay y X.25. Su servicio es de alta prioridad y no garantizado.

B.4.4. Capa AAL5

Este AAL ha sido diseñado para utilizarse bajo TCP/IP y está normalizado en la RFC 1577. AAL Tipo 5 es similar a AAL 3/4 con un programa de simplificación de cabecera (*header*) de información. Este AAL asume que los datos son secuenciales desde el usuario final y usa el bit *Payload Type Indicator* (PTI) para indicar la última celda en transmitirse. Ejemplos de este servicio son el clásico IP sobre ATM, Ethernet sobre ATM, SMDS, y emulación LAN (LANE). Su servicio es de baja prioridad y no garantizado.

Apéndice C

PDH-SDH

C.1. Introducción

El desarrollo de los sistemas de transmisión digitales comenzó en la década de los 70, gracias al desarrollo de las técnicas de *Pulse Code Modulation* (PCM).

A principios de los años 80 los equipos comenzaron a ser cada vez más sofisticados y complejos. El mercado empezó a demandar nuevas tecnologías capaces de crecer de forma jerárquica mediante la multiplexación de líneas de baja capacidad en otras mayores. Pronto se alcanzaron altas cotas de multiplexación, en EEUU se llegó a 140Mbps y en Europa a los 565Mbps. El problema surgió con el alto coste del ancho de banda y de los dispositivos. La solución llegó de la mano de una nueva técnica de multiplexación, la cual permitía combinar flujos ligeramente no síncronos, esta tecnología era más barata y fácil de implementar. Se bautizó con el nombre de jerarquía digital plesiócrona o PDH (*Plesiochronous Digital Hierarchy*).

C.2. PDH (*Plesiochronous Digital Hierarchy*)

PDH permite la transmisión de varios canales de información que están fluyendo a la misma velocidad nominal, pero cuyas tasas binarias pueden variar en torno a un valor nominal.

Las versiones europeas y americanas de PDH difieren ligeramente unas de otras en su forma de funcionamiento, aunque los principios son los mismos. En la versión europea el flujo básico de transmisión se corresponde con 2,048Mbps. Para transmisiones de voz estos 2,048Mbps se segmentan en 30 subflujos de 64kbps para transportar voz más otros dos subflujos de 64kbps dedicados a señalización. Para transmisiones de datos existen otras configuraciones alternativas.

Un reloj interno en los equipos controla la generación de los datos a la tasa de 2Mbps. Esta tasa se permite fluctuar un pequeño porcentaje: diferentes flujos pueden estar (seguramente) corriendo a tasas ligeramente diferentes unas de otras

Para crecer en la jerarquía se multiplexan las tramas de 4 en 4, se selecciona el bit 1 de la primera, luego el bit 1 de la segunda y así hasta la cuarta repitiendo el proceso con todo el flujo de 2048kbps. Además es preciso añadir bits extra para poder demultiplexar la trama: estos bits extra se denominan «justification» o «*stuffing*» bits. Cada estado de multiplexación se denomina nivel PDH, empezando en el nivel 1. En Europa el nivel básico se denomina E1 y se define hasta el nivel E5 (ver cuadro C.1).

Nivel	Tasa binaria
E1	2 Mbps
E2	8 Mbps
E3	34 Mbps
E4	140 Mbps
E5	565 Mbps

Table C.1: Tasas binarias para los niveles PDH europeos

Nivel	Tasa binaria
STM-0	51.84 Mbps
STM-1	155.52 Mbps
STM-4	622.08 Mbps
STM-16	2488.32 Mbps
STM-64	10 Gbps

Table C.2: Tasas binarias SDH

C.3. SDH (*Synchronous Digital Hierarchy*)

En 1980 la *American National Standards Institute* (ANSI) introdujo un estándar para redes de transmisión ópticas para el nombre de «SONET» (*Standard for Optical Networks*). En 1988 la ITU modificó el estándar SONET e introdujo el *Synchronous Digital Hierarchy* (SDH).

SDH ha sido diseñado para ser compatible con PDH. Todos los elementos de la red utilizan la misma señal de reloj lo que permite multiplexar a mayores tasas binarias.

C.3.1. Trama STM-1

Debido a la alta capacidad que ofrece SDH, en la práctica la gran mayoría de los enlaces troncales en redes de telefonía móvil emplearán tramas STM-1. Esta trama se muestra en la figura C.1. Se compone de 270 columnas y 9 filas, lo que suman 2430 Bytes ó 19440 bits. La trama se repite 8000 veces por segundo originando el flujo de 155.52Mbps.

En la figura C.1 se observan que las 9 primeras columnas (81 Bytes) se reservan para tareas de señalización y administración. Por tanto, la carga útil se reduce a 2349 Bytes/trama o 150.336 Mbps.

Como se ha mencionado, SDH es compatible con PDH y, por tanto, una trama STM-1 podrá

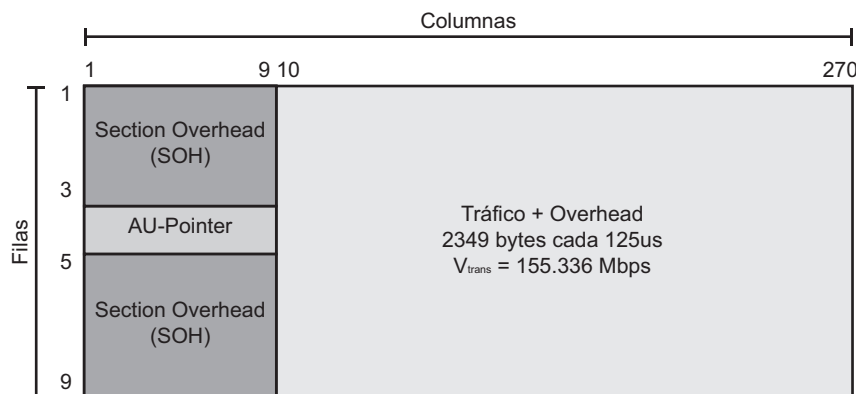


Figure C.1: Estructura de trama STM-1

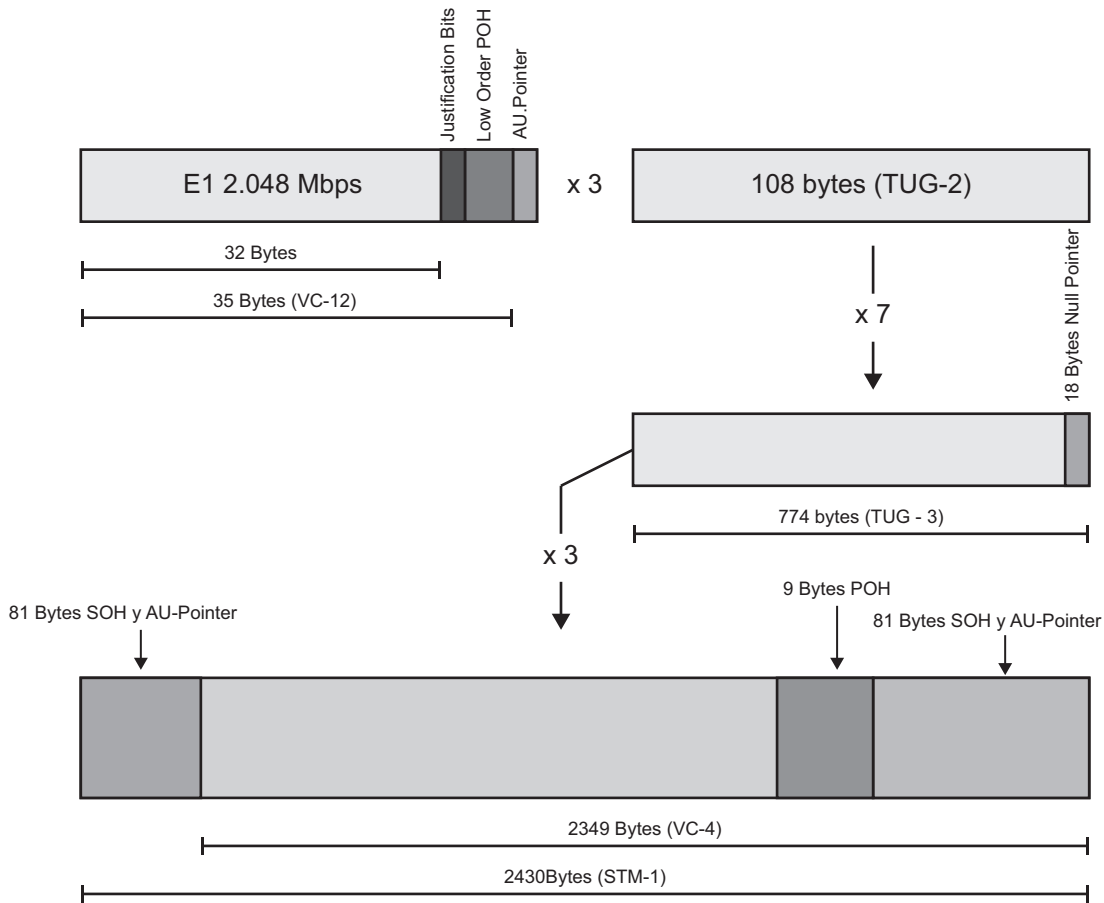


Figure C.2: Mapeado de E1s sobre STM-1

transportar varias tramas de E1s. Cada STM-1 podrá albergar hasta 63 E1s, como se ilustra en la figura C.2. Cada E1 a 2048 kbps producirá $\frac{2.048 \times 10^6}{8000} = 32$ Bytes a transportar en cada trama STM-1. Cada 32 Bytes se incluirán bits de señalización POH, lo que forma un contenedor VC12 de 35 Bytes. A cada VC12 se le añade un bit de administración para formar un *Tributary Unit* 12 (TU-12). Tres TU-12 se combinan para formar un *Tributary Unit Group* (TUG-2) de 108 Bytes. 7 TUG-2 se combinan con un puntero *null* de 18 Bytes para formar un *Tributary Unit Group* de nivel 3 (TUG-3) de 774 Bytes. Finalmente, tres TUG-3 y 9 bytes de señalización más 18 Bytes de «*stuffing*» se combinan para formar un contenedor de nivel 4, VC4, que, junto con los 81 Bytes de señalización y la unidad de administración (AU), forman la trama STM-1.

Apéndice D

Cálculo del ancho de banda medio usado en el RAB Iub ATM

D.0.2. Ancho de banda medio usado en el RAB Iub ATM

Todo el tráfico es cursado por el canal de transporte dedicado o *Dedicated transport Channel* (DCH). La estructura de este canal se muestra en la figura D.1.

En el enlace ascendente o *uplink* (UL) la trama del DCH estima el factor de calidad y añade un CRC a cada TB incluido en la trama DCH.

A la hora de dimensionar, se descartarán los 3 bytes opcionales del campo «*Spare Extension*». Y además los dos últimos bytes de la trama DCH se construyen a partir de un código CRC de 16 bits. Se considera además que el UL DCH FP añade 5 + 4 bytes de cabeceras mientras que el DL DCH FP añade 5 + 2 bytes.

D.0.2.1. Señalización sobre ATM

Cada usuario activo tiene una señalización asociada denominada *Signaling Radio Bearer* (SRB). Durante el intervalo de transmisión de 40ms, *Transmission Time Interval* (TTI), de este SRB, se transmiten 136 ó 128 bits de información de señalización del protocolo RRC. Para transportar estos bits se añaden las cabeceras MAC (*Media Access Control*), FP (*Framing Protocol*), AAL2 y el resto de las cabeceras ATM.

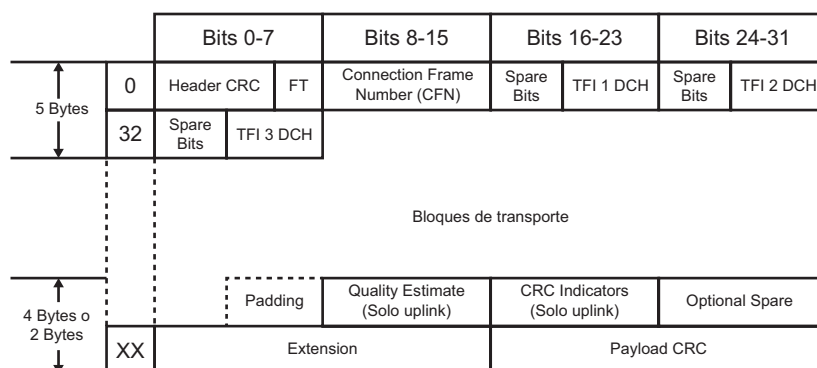


Figura D.1: Estructura de la trama DCH

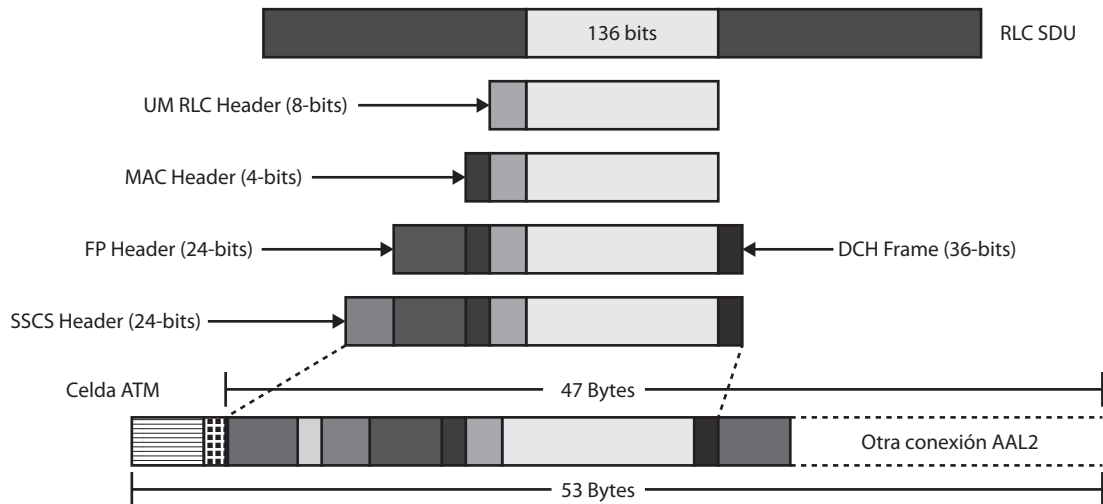


Figura D.2: SRB ATM UL Ancho de banda medio.

Por tanto, el tráfico de señalización en el UL para mensajes RRC supone enviar un segmento de 17 bytes correspondiente al RLC SDU. A éste se añade un byte para formar la RLC PDU. La capa MAC añade 4bits (0.5bytes) para crear la MAC PDU o TB. Puesto que solo hay un *Transport Format Indicator* requerido para el SRB la cabecera FP añade solo 3 bytes. Sin embargo, ya que el TB no es múltiplo par, es necesario completar el último byte («padding»). Los 4.5 bytes que contienen los 4 bits del *padding*, la estimación de calidad, los indicadores de CRC y el CRC se añaden a la carga para crear la trama DCH. El proceso añade $3 + 4.5$ bytes de *overhead*.

A la trama DCH se deben añadir 3 bytes de *Service Specific Convergence Sub-layer* (SSCS) recogidos en el estándar AAL para crear la *Common Part Sub-layer* (CPS) *Service Data Unit* (SDU). En total $17 + 1 + 0,5 + 3 + 4,5 + 3 = 29$ bytes que se envían cada 40 ms, lo que hace una tasa binaria de 5,8 kbps. Ahora bien, en las documentaciones de los equipos Ericsson se recomienda establecer un parámetro de actividad de envío de tramas del 20% haciendo que la tasa baje hasta 1,16 kbps.

En este punto, la tasa de 1.16kbps deberá ser combinada con otras conexiones AAL2. La capa CPS añadirá otro byte de cabecera para crear la carga útil ATM a la que se añaden los 5 bytes de cabecera y formar así la celda de 53 bytes estándar. Para realizar este último paso se multiplica la tasa de bits por el factor $\frac{53}{47}$ obteniendo el resultado final de 1.31kbps. A continuación se muestra a modo de resumen los cálculos explicados y en la figura D.2 se muestra de manera gráfica el proceso de encapsulación.

$$17 + 1 + 0,5 + 3 + 4,5 + 3 = 29\text{bytes} \quad \text{en } 40\text{ms} = 5,8\text{kbps}$$

$$20\%\text{factor actividad} \Rightarrow 1,16\text{kbps}$$

Se suman por último las cabeceras ATM:

$$1,16 \frac{53}{47} = 1,31\text{kbps}$$

Para el caso de la señalización *downlink* se sigue un proceso similar pero teniendo en cuenta que los campos *Quality Estimate* y *CRC Indicators* no son añadidos al DCH. Esto significa que la tasa será ligeramente menor y que en total serán enviados $17 + 1 + 0,5 + 3 + 2,5 + 3 = 27$ Bytes cada

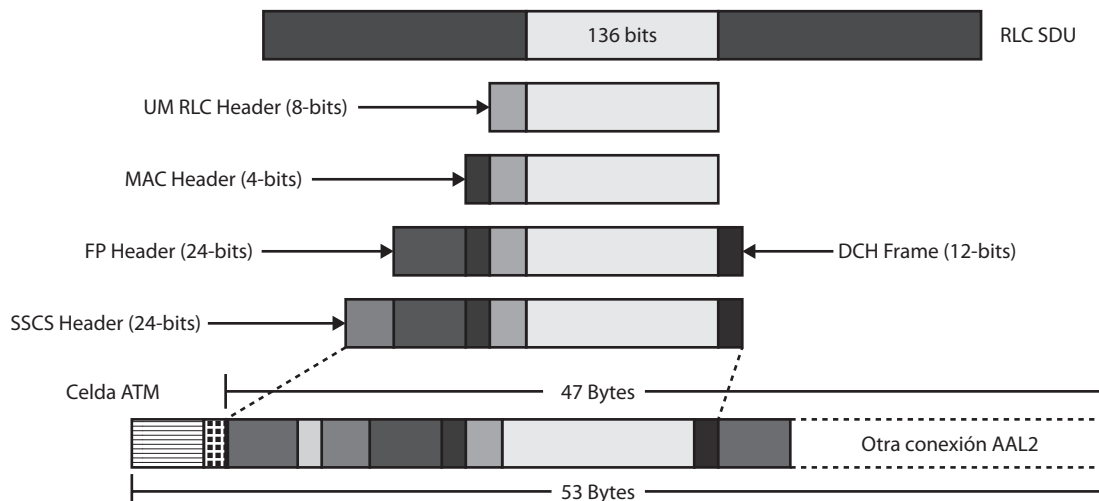


Figura D.3: SRB ATM DL Ancho de banda medio.

40ms. Esto equivale a una tasa de 5,4kbps, a la que, como en el caso del *uplink*, se le aplica el factor de actividad del 20 %, lo que deja una tasa final de 1,08kbps al nivel CPS SDU, que, como antes, se convierte al nivel ATM multiplicando por el factor $\frac{53}{47}$ y obteniendo un resultado de 1,22kbps.

En la figura D.3 se muestra el proceso de encapsulación para el donwlink.

$$17 + 1 + 0,5 + 3 + 2,5 + 3 = 27\text{bytes} \quad \text{en } 40\text{ms} = 5,4\text{kbps}$$

$$20\% \text{factor actividad} \Rightarrow 1,08\text{kbps}$$

Se suman por último las cabeceras ATM:

$$1,08 \frac{53}{47} = 1,22\text{kbps}$$

D.0.2.2. Tráfico de voz sobre ATM

La voz en UMTS puede ser transportada en varios formatos dependiendo del tipo de codificación empleada. Habitualmente se emplea la *Adaptive Multi Rates* (AMR) a una tasa de 12.2kbps.

El AMR *Vocoder* permite tres clases diferentes de bits:

- ⇒ Clase A: Especifica si se transporta sonido o el descriptor de inserción de silencio o *Silence Insertion Descriptor* (SID).
- ⇒ Clase B: Especifica el «*pitch*» del sonido.
- ⇒ Clase C: Especifica el timbre del sonido.

Durante los 20ms de *Transmission Time Interval* (TTI) en los que la voz esta presente, el AMR *Vocoder* a tasa de 12.2kbps produce 81 bits de clase A, 102 bits de clase B y 60 de clase C. Bajo condiciones normales se asumen que la mitad del tiempo sólo una de las dos partes de la comunicación es la que habla mientras que la otra escucha, es decir, se asume un factor de actividad del 50 %. Durante la mitad del tiempo, tiempo de inactividad, el *vocoder* utiliza transmisiones discontinuas (DTX) para salvaguardar los recursos del sistema.

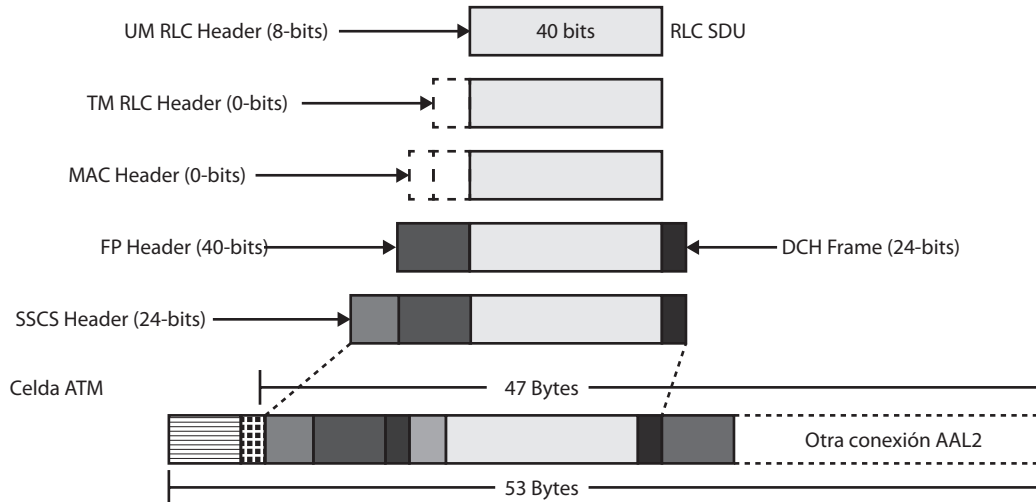


Figura D.4: SID UL ATM Ancho de banda medio, en tramas mudas.

El vocoder también transmite 39 bits del descriptor de inserción de silencio o *Silence Insertion Descriptor* (SID) cada 8 tramas de voz y que se reproducen en el receptor final durante los periodos DTX para crear el ruido de confort. A continuación se calculará el ancho de banda requerido por el SID para el caso del *uplink*.

A los 39 bits se le añade un relleno de 1 bit para formar 40 bits, es decir, 5 bytes y formar una RLC *Service Data Unit* (RLC SDU). En este caso, se opera en modo transparente, por lo que no se omiten cabeceras RLC y tampoco son necesarias las cabeceras MAC. Se añaden 5 bytes de FP y 4 bytes del campo de calidad estimada y los indicadores de CRC, así como el propio CRC, para formar la trama DCH. Por tanto, se han añadido $5 + 4$ bytes de cabecera.

Las cabeceras de 3 bytes de *Service Specific Convergence Sub-layer* (SSCS) se añaden al DCH para formar la *Common Part Sub-layer* (CPS) *ServiceDataUnit* (SDU). En total, $5 + 5 + 4 + 3 = 17$ Bytes (136 bits) a enviar cada 160ms. Esto supone una tasa de 0.85kbps, pero asumiendo el factor de actividad del 50% se reduce hasta 0.425kbps al nivel CPS SDU.

Resta combinar esta tasa binaria con otros flujos AAL2 para formar la carga de 47bytes. La capa CPS añade su byte de cabecera, que se suman a los otros 5 de la cabecera ATM. Por tanto, la tasa binaria final sera $0,425 \frac{53}{47} = 0,5$ kbps. En la figura D.4 se muestra el proceso de ensamblado de las tramas.

Cuando se está transmitiendo voz, entonces el *vocoder* produce 81 bits de clase A, 103 de clase B y 60 de clase C para describir los 20ms de la muestra de voz. Los 81 bits de clase A se rellenan hasta completar 88 bits (11 bytes), los 103 de clase B se convierten en 104 (13 bytes) y los 60 bits de clase C en 64 (8 bytes). La trama inicial se ha convertido entonces a 32 bytes en el nivel RLC SDU. A estos 32 bytes se les añade las cabeceras de *Radio Link Control* (RLC), *Media Access Control* (MAC), *Framing Protocol* (FP), *Service Specific Convergence Sub-layer* (SSCS), *Common Part Sub-layer* (CPS) y ATM como se ha hecho en el caso de no transmitir trama de voz. Se añaden 5 bytes que contienen la calidad estimada, el indicador de CRC y el propio CRC para formar la trama DCH, en total se suman $5 + 4$ bytes. Los 3 bytes de la cabecera *Service Specific Convergence Sub-layer* (SSCS) se añaden al DCH para crear la *Common Part Sub-layer* (CPS) *Service Data Unit* (SDU). En total, $32 + 5 + 4 + 3 = 44$ Bytes (352 bits) que se envían cada 20ms. Esto supone 17.6kbps. Si asumimos el 50% de factor de actividad se reduce hasta 8.8 kbps en el nivel CPS

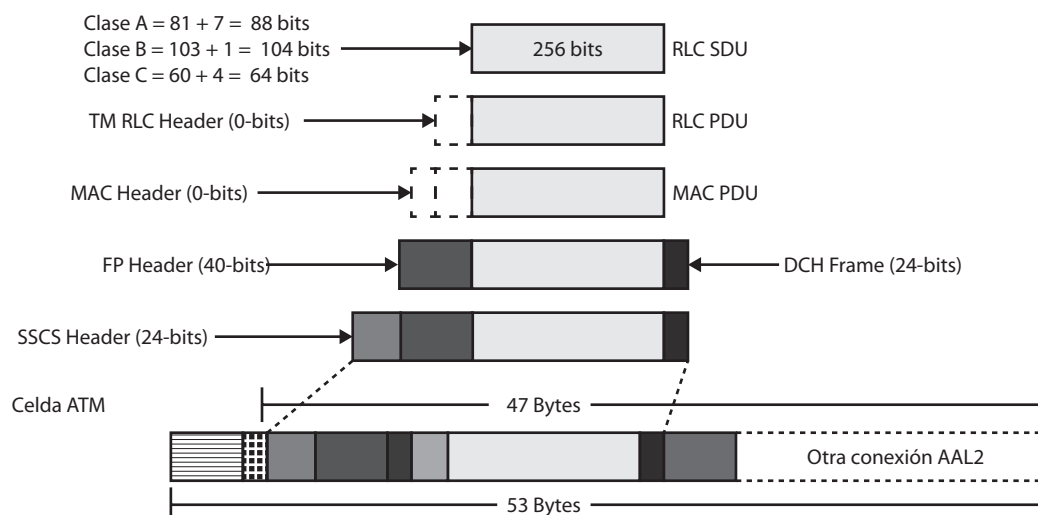


Figura D.5: SID UL ATM Ancho de banda medio, en tramas con muestra de voz.

SDU. Se combinan ahora diversos flujos sobre AAL2, luego se tienen que añadir un byte de CPS y los 5 de la cabecera ATM, en total $8,8 \frac{53}{47} = 9,9$ kbps.

Teniendo en cuenta ambas situaciones, transmisión de voz y transmisión de ruido de confort, se obtiene que la tasa total para el enlace *uplink* supone $9,9 + 0,5 = 10,4$ kbps.

Una vez calculado el enlace ascendente, se hará lo propio con el descendente. Se empieza primero con el caso de estar transmitiendo el ruido de confort para luego calcular la situación de transmisión de tramas de voz.

Para transmitir los 39 bits del SID se necesita añadir uno extra de relleno para completar 5 bytes del protocolo RLC. A nivel MAC, en este caso, no se añaden cabeceras. Puesto que los indicadores CRC no están presentes en el enlace descendente, sólo se necesitan 5 bytes de cabeceras FP y 2 para el CRC para crear la trama DCH. Por tanto, en este proceso se han añadido $5 + 2$ bytes. A continuación, se añaden los 3 bytes del *Service Specific Convergence Sub-layer* (SSCS) al DCH para formar la *Common Part Sub-layer* (CPS) *Service Data Unit* (SDU). En total, $5 + 5 + 2 + 3 = 15$ Bytes (120 bits) a enviar cada 160ms. Esto supone una tasa de de 0.75kbps, que se reduce por el factor de actividad hasta 0.375kbps. Debido a la combinación en AAL2 añadimos el byte de CPS y los 5 de cabecera ATM, en total, obtenemos una tasa de 0.4kbps. En la figura \ref{fig:SIDdl} mostramos el proceso de ensamblado.

En el caso de transmitir voz se utiliza el modo transparente del protocolo RLC para transportar 32 bytes de información. Aquí tampoco se añaden cabeceras MAC. Puesto que los indicadores CRC no están presentes en el enlace descendente, solo se necesitan 5 bytes de cabeceras FP y 2 para el CRC para crear la trama DCH. Por tanto, en este proceso se han añadido $5 + 2$ bytes. A continuación, se añaden los 3 bytes del *Service Specific Convergence Sub-layer* (SSCS) al DCH para formar la *Common Part Sub-layer* (CPS) *Service Data Unit* (SDU). En total, $32 + 5 + 2 + 3 = 42$ Bytes (336 bits) a enviar cada 20ms. Esto supone una tasa de de 16.8kbps, que se reduce por el factor de actividad hasta 8.4kbps. Debido a la combinación en AAL2 se añade el byte de CPS y los 5 de cabecera ATM, en total, obtenemos una tasa de 9.5kbps. En la figura D.7 se muestra el proceso de ensamblado.

Con los dos escenarios contemplados ya puede calcularse el resultado final como la suma de ambas tasas, $0,4 + 9,5 = 9,9$ kbps.

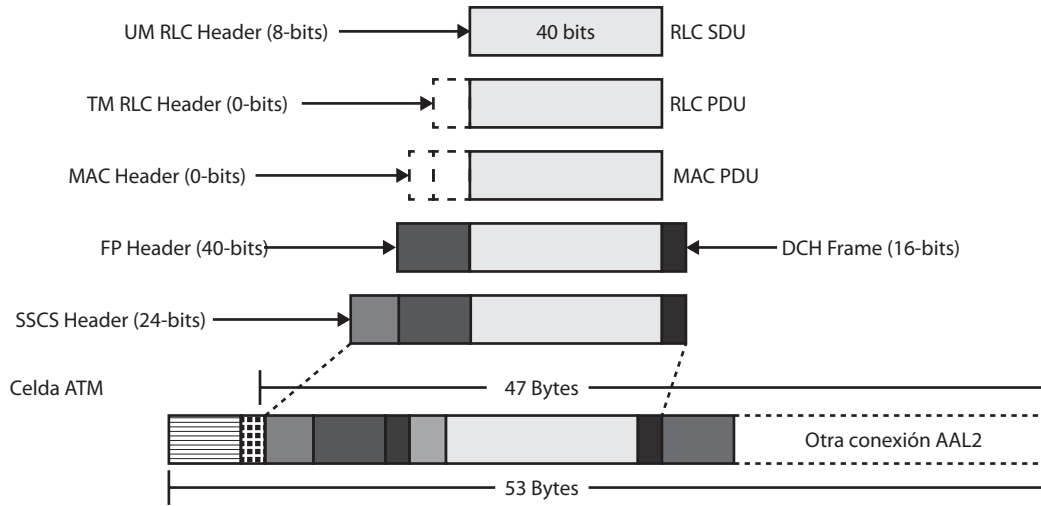


Figura D.6: SID DL ATM Ancho de banda medio, en tramas mudas.

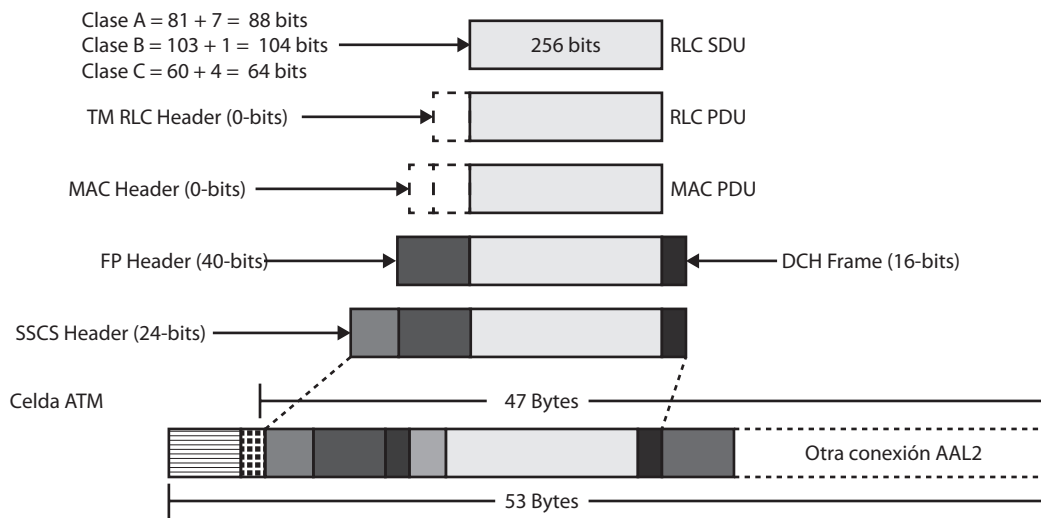


Figura D.7: SID DL ATM Ancho de banda medio, en tramas con muestra de voz.

Tasa	BW medio en DL [kbps]	BW medio en UL [kbps]
AMR 12.2	9.9	10.4
AMR 7.95	7.4	7.9
AMR 5.9	6.1	6.6
AMR 4.75	5.6	6.1
CS 64 kbps	79.7	80.7
CS 57.6 kbps	70.8	71.3
PS 64 kbps	62.6	63.3
PS 128 kbps	123.5	124.2
PS 384 kbps	368.8	370.1
SRB	1.22	1.31

Cuadro D.1: Anchos de banda medio en el canal Iub

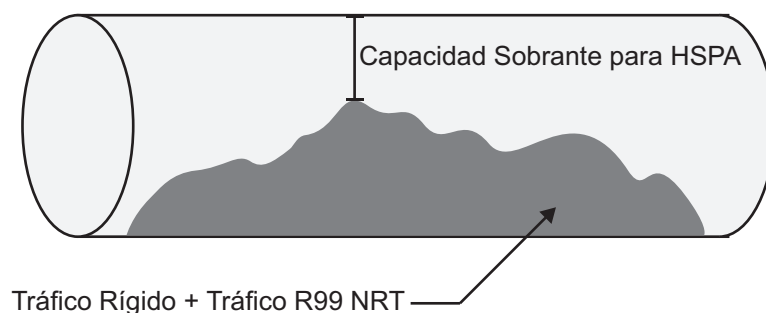


Figura D.8: Tráfico sobrante destinado a HSPA

Lo que se ha hecho para el caso de utilizar como codec AMR a 12.2kbps se puede hacer de forma análoga para el resto de casos, en la tabla D.1 se hace un resumen del resto de situaciones obtenidas aplicando la misma metodología que la empleada en el cálculo de la situación de AMR 12.2kbps.

D.0.3. Algoritmo tráfico HSPA

Puesto que se conocen los erlangs de tráfico de cada servicio rígido y se ha calculado el tráfico medio que demanda en el Iub (tráfico rígido garantizado + tráfico de datos R99 NRT, ver figura D.8), sabiendo el número medio de usuario HSPA (dato de entrada al modelo) puede obtenerse si los usuarios disfrutarán de la máxima velocidad teórica posible, limitada por la interfaz radio, teniendo en cuenta como factor limitante el canal Iub.

Por tanto, el GoS que experimenta cada usuario viene dado por la fórmula (D.1).

$$\text{GoS} = \frac{\text{Capacidad Total} - \sum \text{Throughput}_{\text{Servicio}_i} \text{Erlangs}_{\text{Servicio}_i}}{\text{Usuarios HSPA}} \quad (\text{D.1})$$

Con el valor obtenido en (D.1) se obtendrá el *throughput* que experimenta cada usuario cuando la política es repartir el ancho de banda equitativamente entre todos los usuarios.

Apéndice E

Arquitectura de protocolos de UTRAN/E-UTRAN

E.1. Introducción

En la figura E.1 se muestra una visión global de la arquitectura de protocolos de UTRAN, encuadrada dentro del marco de las comunicaciones que se desarrollan entre los equipos de usuario y el núcleo de red.

Básicamente, un UE puede intercambiar con la red tres tipos de tráfico: tráfico de señalización móvil-red, tráfico de usuario modo circuito (por ejemplo, tramas de voz AMR o datos modo circuito) y tráfico de usuario en modo paquete (por ejemplo, paquetes IP o tramas PPP). Estos tres flujos de información se desarrollan entre el equipo de usuario y el nodo de acceso al núcleo de red correspondiente (MSC o SGSN, según el dominio que se trate). Para alcanzar el destino, es necesario pasar a través de la interfaz radio, la red de acceso y la interfaz Iu, haciendo uso de los servicios que proporciona el estrato de acceso.

Los protocolos de UTRAN se pueden descomponer en dos grandes grupos: los protocolos radio y los protocolos de las interfaces Iub, Iur e Iu. Sobre los primeros, cabe señalar la existencia de una pequeña parte de procedimientos que terminan en el Nodo B, siendo el resto terminados en el RNC. En este segundo caso, el transporte de las tramas radio hasta el RNC se lleva a cabo mediante los servicios que ofrecen los protocolos de la interfaz Iub (así como los de la Iur, en caso de trasposos suaves entre RNC). Finalmente, el tercer eslabón del estrato de acceso corresponde a los protocolos de la interfaz Iu, completándose de este modo el trayecto entre el móvil y el núcleo de red.

E.2. Arquitectura en capas

La figura E.2 muestra la arquitectura de protocolos general de las interfaces Iu, Iur e Iub.

El punto de partida, común para las tres interfaces, es la definición de dos capas de protocolos: la capa de red radio, RNL (*Radio Network Layer*) y la capa de red de transporte, TNL (*Transport Network Layer*). Esta descomposición tiene como objetivo aislar las funciones que son específicas del sistema UMTS, de aquellas otras que dependen de la tecnología de transporte utilizada y que, en su día, podrían ser reemplazadas. Estos dos tipos de funciones se encuadran, respectivamente,

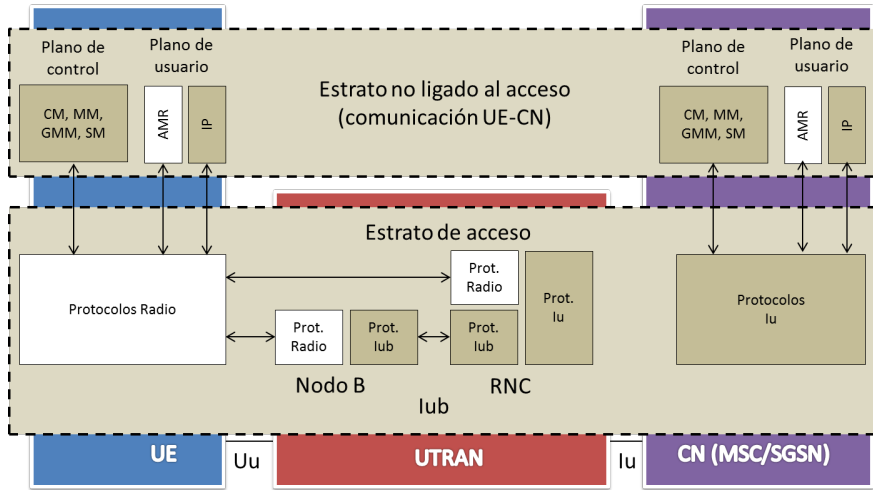


Figura E.1: Arquitectura de protocolos de UTRAN

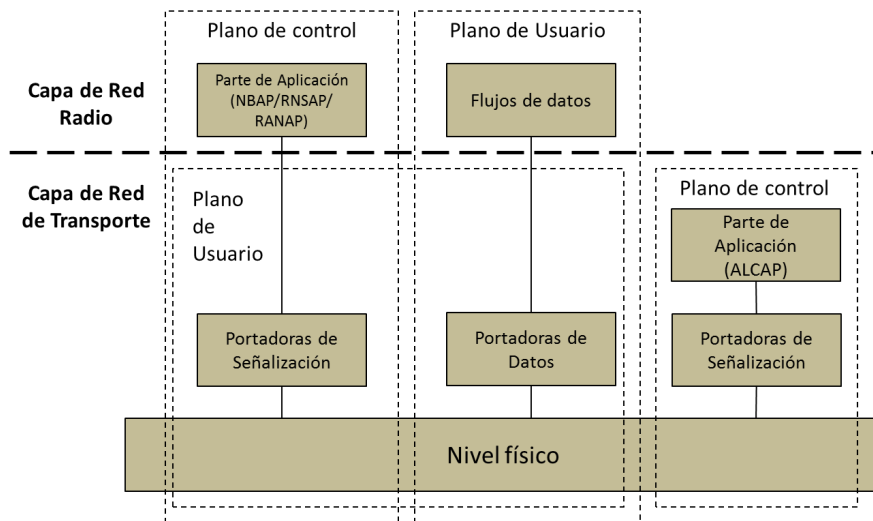


Figura E.2: Arquitectura general de protocolos de las interfaces UTRAN

en las dos capas mencionadas.

En la *Release 99* de UMTS, la capa TNL está basada en ATM, habiéndose seleccionado esta tecnología por ser una de las más flexibles y maduras (al menos en la fecha que se tomó la decisión) para el desarrollo de redes multiservicio con QoS. En cualquier caso, gracias a la descomposición en capas de UTRAN, se facilita el reemplazo en el futuro de ATM por otra tecnología. De hecho, uno de los objetivos de la *Release 6* fue la elaboración de una solución alternativa para UTRAN conocida como IP-RAN, una red de acceso radio UMTS basada en tecnología IP.

Dentro de la capa RNL se distinguen dos planos de comunicación tradicionales: el plano de control y el plano de usuario. El plano de control hace referencia al intercambio de señalización entre los elementos ubicados a cada extremo de la interfaz. Las especificaciones de UTRAN los protocolos de señalización (partes de aplicación) específicos a utilizar en cada interfaz: NBAP para Iub (*Node B Application Part*), RANAP (*Radio Access Network Application Protocol*) para Iu, y RNSAP (*Radio Network Subsystem Application Part*) para Iur.

Por su parte, el plano de usuario hace referencia a los flujos de información que, desde el punto de vista de la interfaz, son meros datos a transportar. Así, por ejemplo, en la interfaz Iub el plano de usuario lo componen los distintos flujos de datos intercambiados sobre la interfaz radio (señalización usuario-red, datos modo circuito y datos modo paquete).

El transporte de los flujos de control y de usuario de la capa RNL se lleva a cabo, respectivamente, mediante las portadoras de señalización y de datos, ubicadas dentro de la capa TNL. Desde el punto de vista de ésta última, toda la información que le suministra la capa RNL es información de usuario.

En función de la interfaz concreta considerada, las portadoras de datos y de señalización se materializan de manera diferente. En este último caso, además, algunas interfaces admiten dos pilas de protocolos alternativas: una pila basada en protocolos del ITU-T, otra basada en las especificaciones del grupo de trabajo SIGTRAN del IETF. Esta segunda opción está encaminada a facilitar la evolución hacia IP de UTRAN.

Como se verá, en ocasiones es necesario un diálogo de señalización entre los extremos de la interfaz para el establecimiento (y liberación) dinámico de las portadoras de datos. Esto explica la presencia de un plano de control dentro de la capa TNL, donde se incluye una parte de aplicación denominada ALCAP (*Access Link Control Application Part*), cuyo significado se comprenderá más adelante.

E.3. Protocolos de la interfaz Iub

En la figura E.3 se muestra la arquitectura de protocolos normalizada para la interfaz Iub.

En la capa de red radio, se consideran dos tipos de información a intercambiar: la información de usuario, correspondiente a los distintos canales radio a transportar sobre Iub, y la información de control, basada en el protocolo NBAP (*Node B Application Part*).

El protocolo NBAP, puede considerarse el equivalente al protocolo BTSM (BTS Management) de GSM. Se trata de señalización intercambiada entre el RNC y el Nodo B. Los procedimientos de señalización NBAP se dividen en dos categorías: procedimientos comunes (C-NBAP) y procedimientos dedicados (D-NBAP). Los primeros corresponden a funciones generales que aplican al nodo B (a la célula UMTS) en su conjunto, mientras que los segundos son procedimientos relativos a un equipo de usuario concreto.

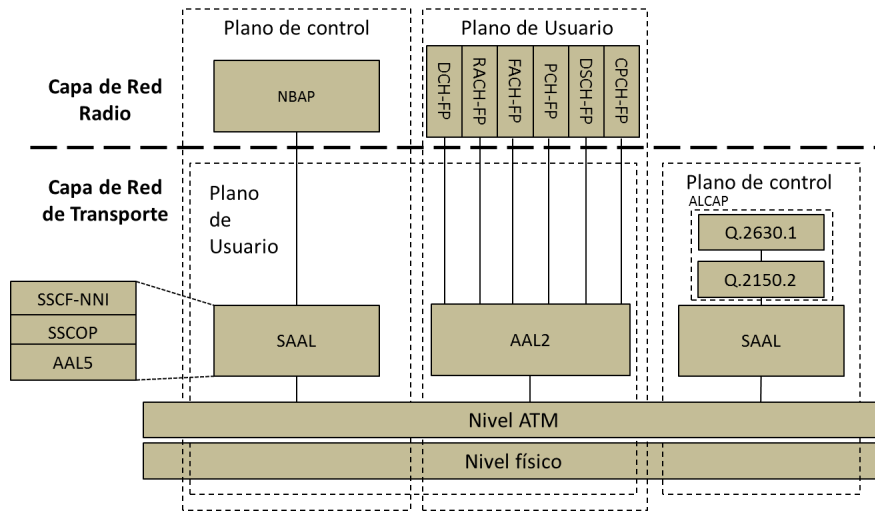


Figura E.3: Arquitectura de protocolos de la interfaz Iub

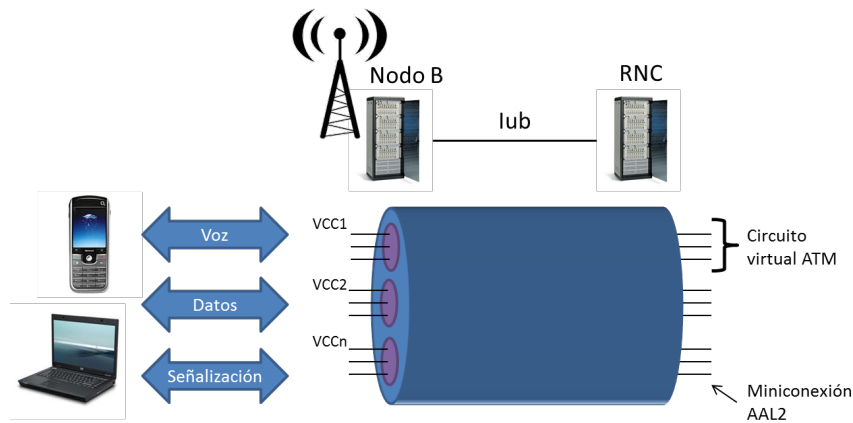


Figura E.4: Transporte de tramas radio sobre la interfaz Iub

La portadora de señalización normalizada para el transporte de NBAP se basa en el protocolo de adaptación ATM para señalización, SAAL (Signalling ATM Adaptation Layer). Este protocolo es en realidad el protocolo AAL5, junto con un par de extensiones que le aportan la fiabilidad que requiere el transporte de señalización: SSCF-NI (Service Specific Coordination Function for support of signalling at the NNI), y SSCOP (Service Specific Connection Oriented Protocol).

Como se ha indicado, los canales radio se prolongan sobre la interfaz Iub hasta alcanzar el RNC. Dicha prolongación se efectúa asociando una conexión virtual (miniconexión) a cada canal radio tal como se indica en la figura E.4. El multiplexado de las miniconexiones sobre la interfaz Iub se efectúa a dos niveles: a nivel de capa de adaptación ATM (AAL) y a nivel ATM.

Las portadoras de datos de la interfaz Iub se basan en el empleo del protocolo AAL2, cuyo funcionamiento se ilustra en la figura E.5.

El protocolo AAL2 permite la multiplexión eficiente de varios flujos de datos (hasta 248) sobre un mismo circuito virtual ATM. En un primer paso, cada flujo de datos se convierte en un flujo de paquetes, vulgarmente denominados minicélulas, con una cabecera de tres octetos y una carga útil de 1 a 45 (ó 64) octetos (el tamaño que más convenga en cada caso). Los flujos de minicélulas resultantes son multiplexados (entrelazados) a continuación formando bloques de 47 bytes. Aña-

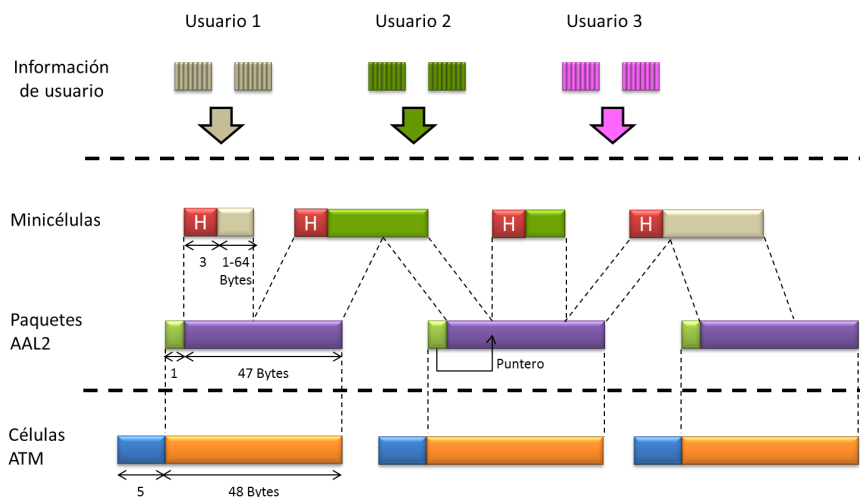


Figura E.5: Funcionamiento del protocolo AAL2

diendo a cada bloque un byte adicional (un puntero que facilita la recuperación ante pérdidas en recepción), se forman los segmentos de 48 octetos con los que se rellena la carga útil de las células ATM de la conexión.

El protocolo AAL2 permite el uso eficiente de una conexión AAL2 al no desperdiciar la capacidad de la carga útil de las células ATM. Si una minicélula no es capaz de llenar la carga útil, la capacidad sobrante se aprovecha para llevar otra minicélula (o al menos parte de ella). Esta cualidad es especialmente interesante en el caso de que los flujos de datos a transportar estén constituidos por bloques de información pequeños, y no se desee agruparlos en bloques mayores con objeto de no aumentar el retardo de empaquetado.

El uso del protocolo AAL2 permite la multiplexión eficiente de hasta 248 canales radio (miniconexiones AAL2) sobre un mismo canal virtual ATM. Puesto que cada canal radio (común o deducado, de usuario o de control) requiere su propia miniconexión AAL2 radio, esta limitación puede dar lugar a la necesidad de definir varios canales virtuales entre cada Nodo B y su RNC. El empleo de varios canales virtuales puede venir dado también por el deseo de separar los distintos tipos de tráfico, con sus distintos requisitos de QoS.

Cada miniconexión AAL2 se identifica de manera única mediante una etiqueta, CID (*Connection Identifier*) que se antepone a cada trama radio a transportar. La negociación del valor de CID a utilizar para cada miniconexión se lleva a cabo mediante el protocolo de señalización AAL2, al que las especificaciones de UMTS hacen referencia dentro del protocolo ALCAP (*Access Link Control Application Part*). En la figura E.6 se ilustra el uso de dicho protocolo para el establecimiento y liberación de una miniconexión AAL2.

La recomendación Q.2630.1 no establece cómo han de transportarse los mensajes de señalización AAL2, permitiéndose utilizar el mecanismo más adecuado en cada caso. En el caso de Iub, la especificación ALCAP establece, al igual que en el caso de NBAP, el empleo de SAAL. El transporte de señalización AAL2 sobre SAAL se efectúa de acuerdo a la recomendación Q.2150.2 del ITU-T.

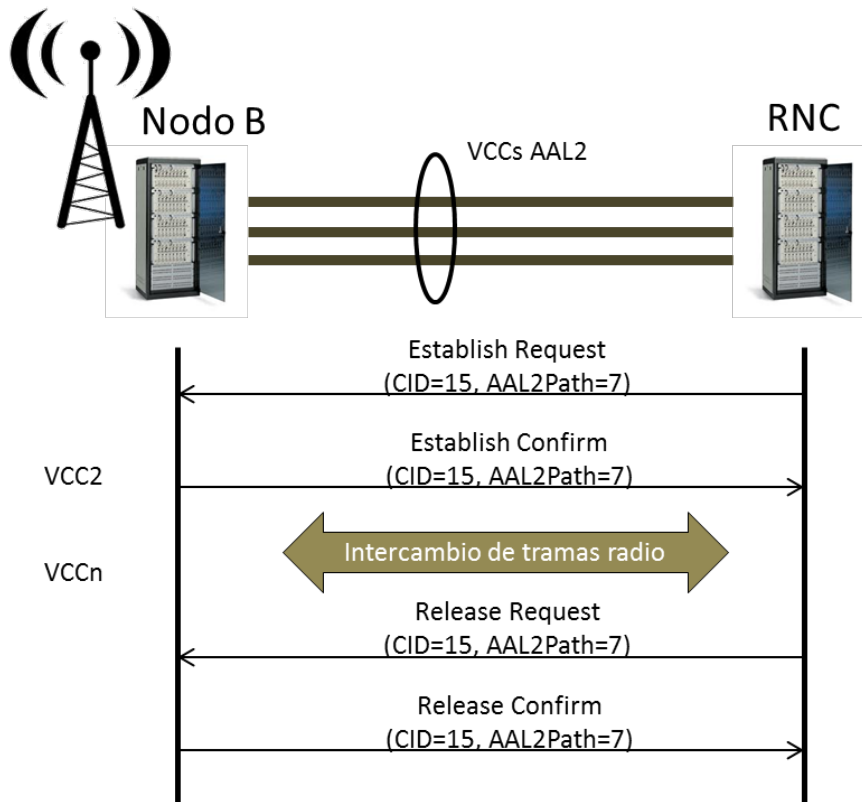


Figura E.6: Establecimiento y liberación de miniconexiones AAL2

E.4. Protocolos de la interfaz Iur

Como se sabe, los traspasos suaves (soft-HO) en UMTS permiten que en un momento dado el móvil pueda comunicarse simultáneamente con varios nodos B. Si los nodos B pertenecen al mismo RNC, la diversidad requiere la utilización de tantas miniconexiones AAL2 como nodos B se estén utilizando, tal como se muestra en la parte izquierda de la figura E.7. En el caso de que los nodos B pertenezcan a distintos RNC, es necesario habilitar también un trayecto entre los D-RNC y el S-RNC, a través de la correspondiente interfaz Iur, tal como se representa en la parte derecha de la misma figura. El transporte de las tramas radio sobre la interfaz Iur se efectúa, al igual que en la interfaz Iub, mediante miniconexiones AAL2.

En la figura E.8 se muestra la arquitectura de protocolos normalizada para la interfaz Iur. El

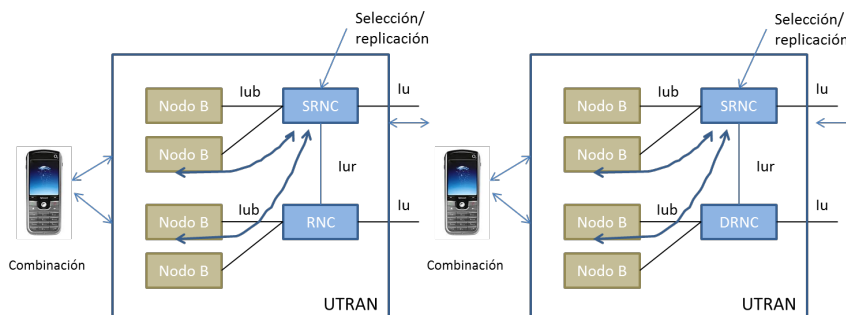


Figura E.7: Impacto de los traspasos suaves (Soft-HO) en UTRAN

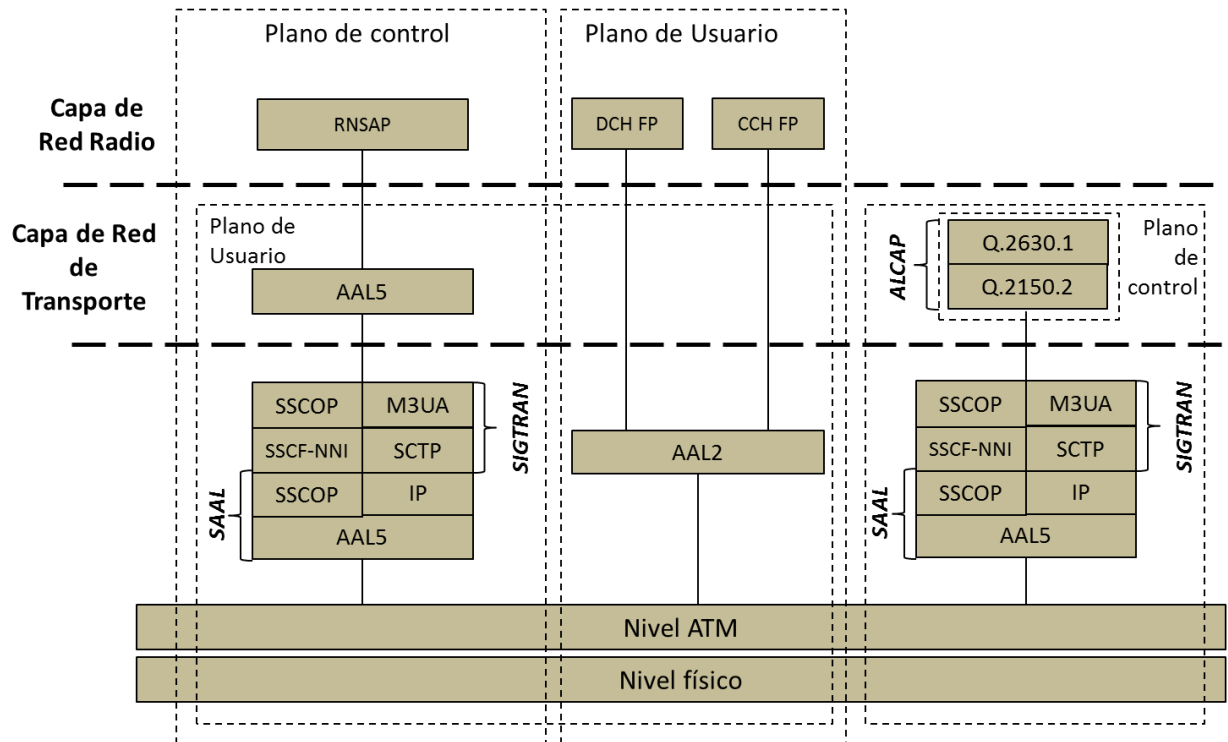


Figura E.8: Arquitectura de protocolos de la interfaz Iur

intercambio de señalización entre D-RNC y S-RNC se efectúa mediante la parte de aplicación RNSAP (Radio Network Subsystem Application Part).

Los mensajes RNSAP se transportan sobre conexiones SCCP (Signalling Connection Control Part), uno de los bloques normalizados de la arquitectura de protocolos SS7, contemplándose dos opciones para la realización de las portadoras de señalización: la solución conforme al modelo ITU-T (SSCP/MTP3b/SAAL), y la basada en los protocolos SIGTRAN del IETF (SSCP/M3UA/SCTP/IP/AAL5). La segunda opción tiene como objetivo facilitar la migración hacia IP de la TNL.

La información de usuario a transportar sobre la interfaz Iur son las tramas radio (datos y señalización móvil-red) intercambiadas entre los RNC durante la situación de traspaso suave. Al igual que en la interfaz Iub, dichas tramas se encapsulan sobre el protocolo FP (Framing Protocol), siendo transportados mediante portadoras de datos basadas en miniconexiones AAL2.

El transporte de señalización AAL2 en este caso asume el empleo del bloque MTP3 del SS7 lo que requiere el adaptador definido en la recomendación Q.2150.1. al igual que para RNSAP, las portadoras de señalización para ALCAP admiten las opciones ITU-T y SIGTRAN.

Los procedimientos de RNSAP están divididos en cuatro grupos: movilidad básica, transporte de canales dedicados (DCH), transporte de canales comunes y procedimientos globales. Dentro de los distintos procedimientos se incluyen los siguientes: establecimiento, reconfiguración, y liberación de enlaces radio UE/DRNC/SRNC; medidas sobre recursos dedicados en DRNS; control de potencia sobre DRNS (sentido descendente); transferencia de señalización de canales de control comunes entre DRNC y SRNC; avisos («Paging») sobre DRNS; reubicación de SRNC; y gestión de errores.

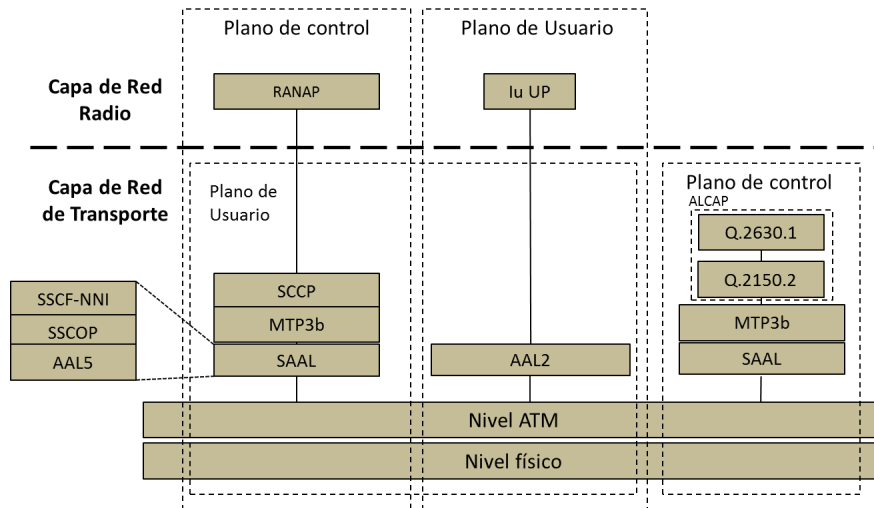


Figura E.9: Arquitectura de protocolos de la interfaz Iu-CS

E.5. Protocolos de la interfaz Iu-CS

En la figura E.9 se muestra la arquitectura de protocolos de la interfaz Iu-CS, mediante el que una RNC se conecta a la MSC de la cual depende.

Los diálogos de señalización entre la RNC y la MSC se llevan a cabo mediante el protocolo RANAP (Radio Access Network Application Part). Al igual que su equivalente en GSM, el protocolo BSSAP, los mensajes RANAP hacen uso de la capa SCCP del SS7. La portadora de señalización utilizada en Iu-CS solamente contempla una opción para el transporte de RANAP: SCCP/MTP3b/SAAL.

Dentro de los procedimientos RANAP cabe distinguir dos grupos. Por un lado están los relativos a los diálogos de señalización desarrollados entre el RNC y la MSC (o el SGSN, ya que como se verá en el siguiente apartado, RANAP también se usa en la interfaz Iu-PS). Dentro de ellos se incluyen los siguientes procedimientos de gestión de la interfaz Iu; asignación, modificación y liberación de RAB; reubicación de SRNC; traspasos «hard»; avisos («paging»); trazado de actividad de un móvil; control de cifrado e integridad; y gestión de informes («reports»).

El otro grupo de procedimientos RANAP tiene que ver con el transporte de señalización móvil-red sobre la interfaz Iu-CS (o Iu-PS) mediante su encapsulado sobre un par de mensajes RANAP definidos al efecto: mensaje inicial (que permite a la red iniciar la conexión de señalización con el usuario sobre Iu) y el mensaje de transferencia directa (para el intercambio de los mensajes subsiguientes).

En el plano de usuario, los datos (voz o datos modo circuito) se encapsulan sobre el protocolo UP User Part Framing Protocol), y se entregan para su transporte sobre miniconexiones AAL2.

El transporte de señalización AAL2 sobre Iu-CS hace uso de MTP3, empleándose el adaptador Q.2150.1, el mismo que en la interfaz Iur. En el caso de Iu-CS, sin embargo, sólo se contempla el transporte de ALCAP sobre una pila ITU-T.

E.6. Protocolos de la interfaz Iu-PS

En la figura E.10 se muestra la arquitectura de protocolos de la interfaz Iu-PS.

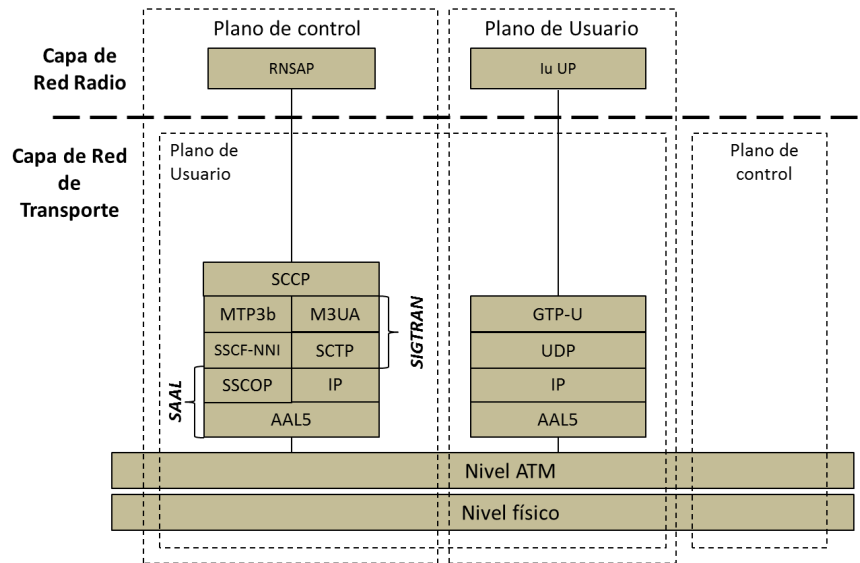


Figura E.10: Arquitectura de protocolos de la interfaz Iu-PS

Los diálogos de señalización entre RNC y SGSN se basan, al igual que en Iu-CS, en el protocolo RANAP. En el caso de Iu-PS, sin embargo, se admiten las opciones ITU-T y SIGTRAN para las portadoras de señalización. La segunda alternativa está encaminada a facilitar la migración de UTRAN a IP.

Las portadoras de datos en la interfaz Iu-PS se realizan mediante túneles IP sobre AAL5, utilizando el protocolo GTP. El establecimiento y liberación de los túneles se lleva a cabo mediante procedimientos RANAP, motivo por el cual el plano de control de la capa TNL se muestra vacío en la figura (nótese que en esta interfaz no se requiere ALCAP al no emplearse portadoras de datos en AAL2). En la figura E.11 se muestra con mayor detalle la secuencia de encapsulado de los paquetes de usuario sobre la interfaz Iu-PS.

Los datos de usuario (un paquete IP, una trama PPP, etc) se encapsulan en paquetes que añaden una cabecera que contiene, entre otras cosas, el identificador de túnel. El paquete GTP se encapsula sobre UDP/IP, entregándose a la capa AAL5 para su segmentación en células ATM. Nótese que, además de su empleo en la interfaz Gn, como en GPRS, en UMTS el protocolo GTP también se utiliza en la interfaz Iu-PS.

La figura E.12 permite comparar las diferencias en el plano de usuario entre GPRS y UMTS (por motivos de claridad, se omiten los protocolos entre estaciones base y controladores).

Por un lado, se encuentran las diferencias debidas al empleo de una nueva interfaz radio, que se traducen en el empleo de nuevos protocolos radio. Esto se aplica a los bloques RLC (Radio Link Control) y MAC (Medium Access Control), distintos a pesar de mantener la misma denominación en ambos sistemas, así como al protocolo PDCP (Packet Data Convergence Protocol), responsable de efectuar la compresión de cabeceras IP, siguiendo un enfoque similar al del bloque SNDCP de GPRS.

La otra gran diferencia se deriva de la utilización de ATM en vez de «Frame Relay» en la interfaz entre la red de acceso y el núcleo de red. La pila de protocolos utilizada sobre Iu-PS, basada en el uso de GTP, permite prescindir de las capas BSSGP y LLC.

A modo de resumen, en la figura E.13 se muestra de manera simplificada cómo se transportan sobre la UTRAN los distintos flujos de información intercambiados entre el móvil y el núcleo de

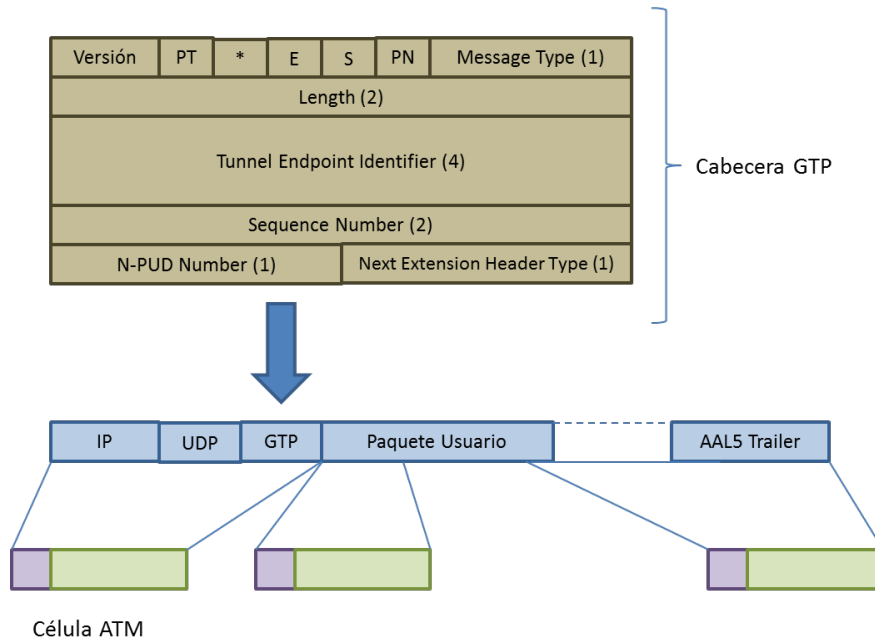


Figura E.11: Encapsulado de paquetes de usuario en la interfaz Iu-PS

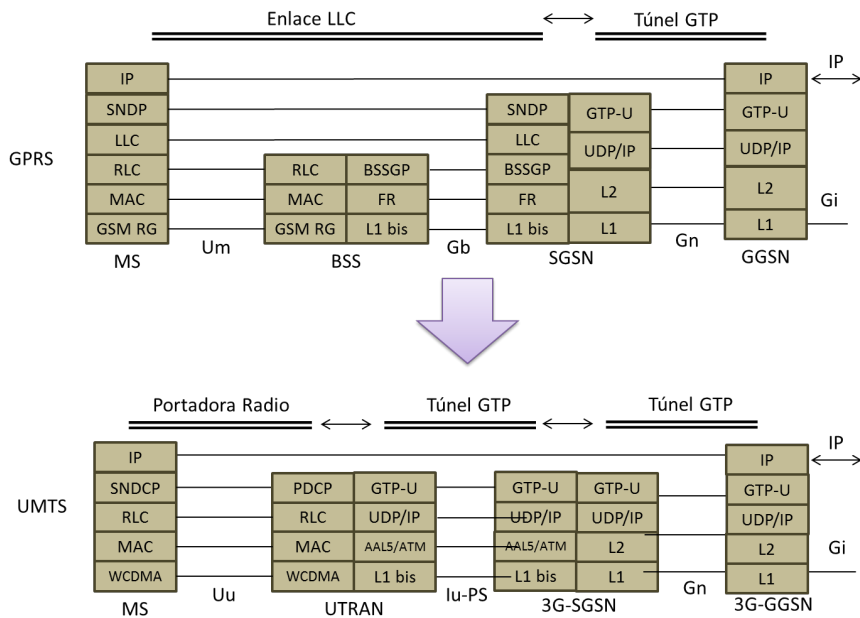


Figura E.12: Comparación de los planos de usuario de GPRS y de UMTS

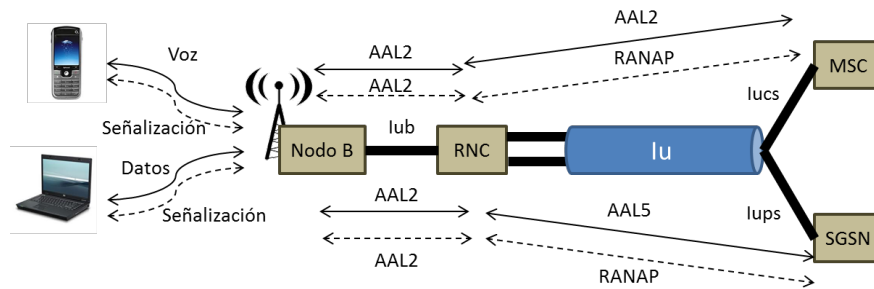


Figura E.13: Transporte de datos y señalización móvil-red sobre UTRAN

red.

La parte superior de la figura corresponde a una comunicación modo circuito entre el móvil y una MSC. En la interfaz Iub, la señalización móvil-red y el tráfico de usuario (por ejemplo, paquetes de voz AMR) se transportan sobre sendas miniconexiones AAL2. Al llegar al interfaz Iu-CS, el tráfico de usuario prosigue hacia la MSC sobre AAL2, mientras que los mensajes de señalización se encapsulan sobre RANAP.

En la parte inferior de la figura se representa una comunicación modo paquete entre un terminal UMTS y un SGSN. La solución sobre la interfaz Iub es la misma que en una comunicación modo circuito, utilizándose sendas miniconexiones AAL2 para la señalización móvil-red y los datos de usuario (por ejemplo, paquetes IP). Al llegar a la interfaz Iu-PS, el tráfico de usuarios se encapsula sobre un túnel IP y se inyecta sobre una conexión AAL5, mientras que la señalización móvil-red se encapsula sobre RANAP.

E.7. Protocolos radio

En la medida que la gran complejidad de las especificaciones de UMTS en el marco del 3GPP ha conducido a la división del trabajo en diferentes grupos TSG (Technical Specification Groups), es lógico que cada uno de los grupos haya desarrollado las arquitecturas de referencia de protocolos que le son propios. En el caso de TSG RAN (Radio Access Network), se define la arquitectura de protocolos de la interfaz radio, tal y como se muestra en la figura E.14 donde se presenta tanto la parte del UE como de UTRAN.

La estructuración de los protocolos sigue los principios de los tres niveles OSI (Open System Interconnection), distinguiéndose entre las capas 1 (nivel físico), capas 2 (subnivel MAC - Medium Access Control - y subnivel RLC - Radio Link Control) y la componente RRC (Radio Resource Control) de la capa 3. En este modelo de referencia se hace distinción entre el plano de usuario y el plano de control. En el primero de ellos se incluye también el protocolo PDCP (Packet Data Convergence Protocol) para la compresión de cabeceras.

Inciendo en la figura E.14, se aprecia la incorporación de los conceptos de SRB (Signalling Radio Bearer), RB (Radio Bearer), canales lógicos y canales de transporte. Por otra parte, las líneas discontinuas indican que RRC tiene capacidad de configurar diferentes parámetros de L1, MAC y RLC. La bidireccionalidad de las flechas indica que las capas bajas trasladan diferentes medidas hacia la capa RRC. En el caso del UE es a través de los mensajes del protocolo RRC que estas medidas se pueden trasladar a UTRAN (en particular, a la RNC). La información que proporcionan los UE se añade a las propias medidas que se realicen en las capas bajas de UTRAN

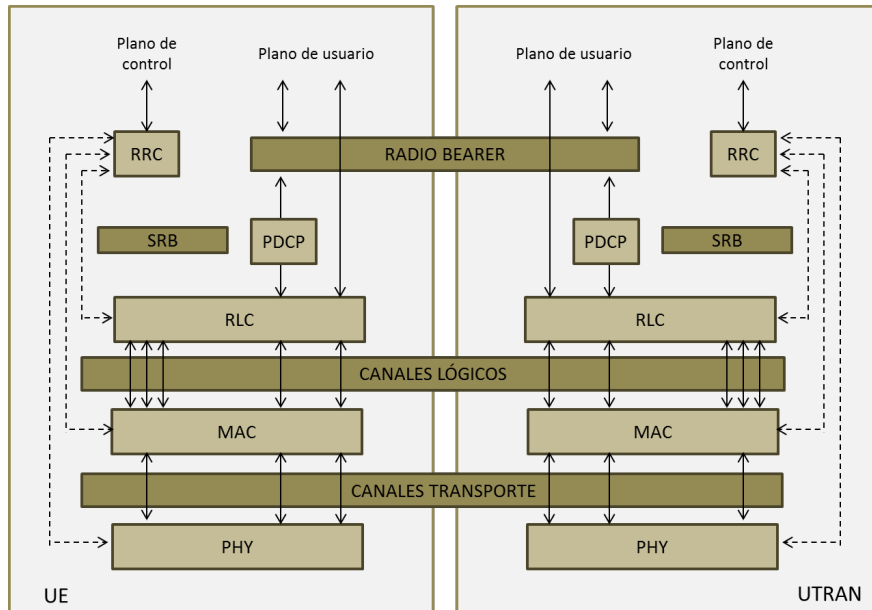


Figura E.14: Arquitectura de protocolos de la interfaz radio

y se trasladan al nivel RRC. Con todo ello, las funcionalidades de control y gestión de los recursos radio RRM (Radio Resource Management), ubicadas en la RNC, se trasladan a los UE a través de mensajes RRC. Esto es, el control local y las medidas locales se tratan a través de los SAP de control entre RRC y las capas bajas, mientras que las transferencias UE-UTRAN se llevan a cabo mediante mensajes RRC propiamente dichos.

En este apartado se incluyen los elementos más destacados de los diferentes protocolos, de manera sintetizada. Dada la extensión y complejidad de especificación de algunos de los protocolos radio, se refiere al lector interesado en mayor nivel de detalle a las especificaciones del 3GPP.

E.7.1. Protocolo RRC (Radio Resource Control)

El protocolo RRC, es fundamental en la gestión de la interfaz radio, ya que soporta las funcionalidades de la gestión de recursos radio. Los mensajes que transporta el protocolo RRC fundamentalmente se corresponden a:

1. Flujos de señalización de capas superiores, a los que el protocolo RRC proporciona el segmento UE-UTRAN de la conexión de señalización a las capas superiores. La extensión del camino UE-UTRAN para llegar a comunicar UE con CN es a través de la interfaz Iu y el protocolo de señalización RANAP.

Por ejemplo, los mensajes que generan las capas superiores del UE se encapsulan en un mensaje del protocolo RRC, que transporta esta señalización sobre la interfaz radio sobre la SRB y llega hasta la RNC. Una vez allí, el mensaje RRC se traduce en un mensaje RANAP, llegando finalmente a través de la interfaz Iu hasta la CN. La figura E.15 muestra cómo a través de la RAB se dispone de un camino de comunicación de datos entre UE y la CN mientras que el camino de señalización se consigue en dos tramos: la SRB sobre protocolo RRC y la Iu SB sobre protocolo RANAP.

2. Mensajes de señalización asociados directamente a la interfaz radio: medidas, configuración,

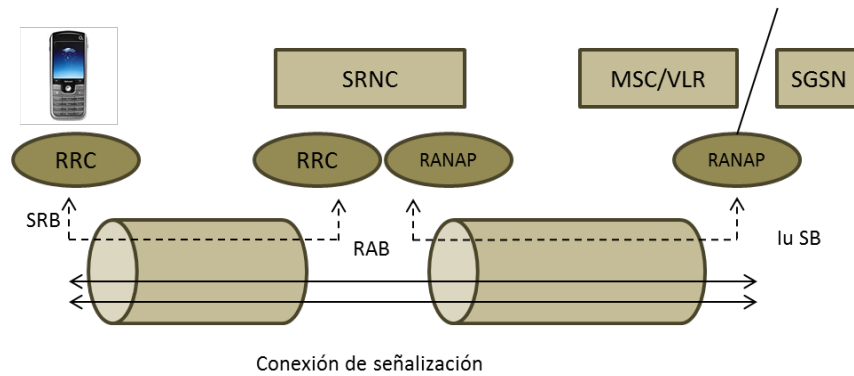


Figura E.15: SRB, RAB y conexión de señalización

Funciones del protocolo RRC
«Broadcast» de información
Establecimiento, mantenimiento y liberación de conexiones RRC entre UE y UTRAN.
Establecimiento, reconfiguración y liberación de Radio Bearers.
Asignación, reconfiguración y liberación de recursos radio para la conexión RRC.
Funciones de movilidad de la conexión RRC.
Control de la QoS demandada.
Reportes de medidas del UE y control de los mismos.
Outer loop power control.
Control de cifrado

Cuadro E.1: Principales funciones RRC

etc, los cuales controlan la asignación de recursos radio al UE, permite al MAC arbitrar entre RB dentro de los recursos radio asignados, etc.

Por ejemplo, se definen mensajes RRC para trasladar medidas que haya realizado el UE hasta la red, donde se podrán evaluar y, en su caso, tomar las decisiones y actuaciones que se crean convenientes. Nótese que, como se muestra en la figura E.15, la capa física del UE realiza medidas (por ejemplo de nivel de potencia recibida de diferentes bases), que localmente se trasladan a la capa 3 (línea discontinua de la figura E.15 entre PHY y RRC). Será desde L3 que se generará el mensaje RRC que se traslada a la red (línea discontinua en la figura entre RRC del UE y RRC de UTRAN) y, una vez allí, se pone a disposición de la entidad RRM. Caso de que esta medida indique la necesidad de alguna actuación sobre la interfaz radio (por ejemplo que el UE debe reducir su velocidad de transmisión) la entidad RRM insta la transmisión de un mensaje RRC al UE (línea discontinua en la figura entre RRC de UTRAN y RRC de UE). Llegado el mensaje al UE e interpretado el contenido del mismo a nivel L3, a través de las interfaces de control local se configurarían las capas más bajas (líneas discontinuas de la figura entre RRC de UE y RLC/MAC/PHY del UE).

Las funciones del protocolo RRC son, como su nombre indica, las asociadas al control de la interfaz radio y que se resumen en la tabla E.1. En cuanto a los servicios que proporciona RRC, quedan recogidos y sintetizados en la tabla E.2.

Servicios proporcionados por el protocolo RRC
«General Control». El GC SAP proporciona el servicio de realizar el «broadcast» de cierta información sobre todos los UE en una cierta área geográfica.
«Notification». El Nt SAP proporciona el servicio de «paging» y notificación sobre una cierta área geográfica pero dedicado a un UE específico.
«Dedicated control». El DC SAP que proporciona la capacidad de establecer y liberar las conexiones y transferir los mensajes haciendo uso de dichas conexiones.

Cuadro E.2: Servicios que proporciona RRC

E.7.1.1. Estados RRC

La descripción del funcionamiento y control de la interfaz radio es sumamente compleja, por lo que se hace indispensable realizar una definición estructurada de los distintos elementos, funciones y procedimientos involucrados. A nivel RRC se definen diferentes estados, y se asocia a cada uno de ellos una serie de características, capacidades y tareas. También se identifican los eventos que originan las transiciones entre estados, y con ello se establece un marco claro de operación a nivel de la interfaz radio. En particular, los dos modos de operación básicos en que puede encontrarse un UE son:

1. «Idle», caracterizado porque no hay una conexión de señalización establecida. En caso de que el móvil desee establecer una conexión RRC puede realizarla mediante el acceso a través del canal de control de la célula en cuestión. Si la red recibe una llamada dirigida a un UE registrado, sabe (en la mayor parte de casos) el área de registro y puede enviar un «paging».
2. «Connected», lo que supone que hay una conexión RRC establecida. Al UE se le asigna una RNTI (*Radio Network Temporary Identity*) para emplearse como identidad en los canales de transporte comunes. El nivel de conexión del UE a UTRAN depende de los requerimientos de QoS de las portadoras radio activas y de las características del tráfico sobre dichas portadoras. Se distingue:
 - a) URA_PCH. En este estado no hay posibilidad de actividad en UL. No puede utilizarse el canal lógico DCCH. Cualquier actividad en el uplink ocasiona la transición a Cell_FACH. Si la red quiere iniciar actividad debe hacer un «paging». La posición del móvil se conoce a nivel de URA (UTRAN Registration Area) cuando el móvil realizó por última vez un «URA Update» en Cell-FACH.
 - b) Cell_PCH. En este estado no hay posibilidad de actividad en UL. No puede utilizarse el canal lógico DCCH. La transición a Cell-FACH se lleva a cabo tras un «paging» o bien cuando se quiere realizar un acceso en el «uplink». En este estado la posición del móvil se conoce a nivel de célula cuando el móvil realizó por última vez un «cell update» en Cell-FACH.
 - c) Cell_FACH. En este estado el UE dispone de DCCH y, si está configurado, un DTCH. Tiene asignado un canal de transporte común en UL (RACH), que lo puede usar en cualquier momento de acuerdo con el procedimiento de acceso del mismo. No tiene un canal físico dedicado. La posición del UE es conocida en UTRAN a nivel de célula según la célula donde el UE realizó la última actualización.

- d) Cell_DCH, dispone de un canal dedicado en «uplink» y «downlink». Se llega desde «Idle» a través del «setup» de una conexión RRC o bien estableciendo un canal dedicado desde Cell_FACH. La posición del móvil se conoce a nivel de celda.

E.7.1.2. Procedimientos RRC

Las distintas capacidades que se tienen en la operación y gestión de la interfaz radio de UMTS se materializan a través de la ejecución de procedimientos, que definen la interrelación necesaria y ordenada entre los diferentes elementos de la red y los protocolos asociados para que efectivamente se puedan llevar a cabo las funcionalidades de la red de manera adecuada.

A continuación se resumen los principales procedimientos RRC.

Procedimientos de gestión de la conexión RRC

- ⇒ «Broadcast» de información de sistema.
- ⇒ «Paging».
- ⇒ Establecimiento de conexión RRC.
- ⇒ Liberación de conexión RRC.

Procedimientos de control de la Radio Bearer

- ⇒ Procedimientos de reconfiguración.
- ⇒ Control de la combinación de «Transport Format».

Procedimientos de movilidad de conexión RRC

- ⇒ Actualización del «Active Set».
- ⇒ «Hard-handover».
- ⇒ «Inter-RAT handover» desde UTRAN.
- ⇒ Reselección de célula inter-RAT a UTRAN.
- ⇒ Cambio de orden de célula Inter-RAT a UTRAN.

Procedimientos de medidas

- ⇒ Intra-frecuenciales.
- ⇒ Inter-frecuenciales.
- ⇒ Inter-RAT.
- ⇒ Volumen de tráfico.
- ⇒ Calidad.
- ⇒ Internas del UE.
- ⇒ Posicionamiento del UE.

Además de los procedimientos de gestión de la conexión RRC y la movilidad asociada o los procedimientos de control de las portadoras radio, en UTRA se otorga una gran importancia a los procedimientos de medidas, ya que constituyen un elemento de soporte esencial para los algoritmos de RRM que, en definitiva, utilizarán estas medidas como información de entrada para la toma de decisión de gestión de la interfaz radio.

UTRAN puede controlar las medidas del UE bien a través de «broadcast» o de mensajes RRC específicos. En los procedimientos de medidas y de reporte de las mismas se deben definir adecuadamente los objetos de medida, la cantidad a medir (y su proceso de filtrado o promediado), el criterio para reportar la medida (por ejemplo, medidas periódicas o disparadas por eventos), etc. El UE debe ser capaz de realizar distintas medidas en paralelo y controlarlas y reportarlas de forma independiente. Las células que se monitorizan se distinguen en tres grupos excluyentes entre sí: a) Células del «active set» (células a las que el UE se encuentra conectado), b) células no incluidas en el «active set» pero que están incluidas en el CELL_INFO_LIST identificadas como «Monitored set» y c) células detectadas por el UE que no están ni en el «active set» ni el «monitored set».

Los procedimientos de medidas pueden ser de «measurement control» (mediante este procedimiento se establece, modifica o cancela una medida determinada en el UE o de «measurement report» (el propósito de este procedimiento es transferir los resultados de las medidas desde el UE a UTRAN). Lógicamente, UTRAN debe tener en cuenta las capacidades del UE antes de solicitarle que pase a realizar ciertas medidas.

E.7.2. RLC

La especificación TS 25.322 del 3GPP describe la arquitectura RLC y sus diferentes entidades desde el punto de vista funcional, pero no especifica ni restringe las implantaciones que se realicen del mismo. Las principales funciones RLC son las siguientes:

- ⇒ Segmentación.
- ⇒ Concatenación.
- ⇒ Transferencia de datos.
- ⇒ Corrección de errores.
- ⇒ Entrega ordenada de PDUs a las capas superiores.
- ⇒ Detección de duplicidades.
- ⇒ Control de flujo.
- ⇒ Comprobación de número de secuencia.
- ⇒ Detección y recuperación de errores de protocolo.
- ⇒ Cifrado.

Por otra parte, los principales servicios asociados al protocolo son los siguientes:

- ⇒ Transferencia de datos en modo transparente.
- ⇒ Transferencia de datos en modo no reconocido: detección de errores y entrega inmediata.

⇒ Transferencia de datos en modo reconocido: entrega única de errores, entrega única y entrega ordenada.

Hay tres tipos de entidades RLC:

1. TM (Transparent Mode), que tendrá una entidad transmisora y otra receptora para cada servicio, y que utilizará un único canal lógico.
 - a) Entidad TM RLC transmisora. Recibe RLC SDU de los niveles superiores, que pueden requerir o no de segmentación para dar lugar a un múltiplo entero de TDM PDU y se someten a la capa inferior a través de los canales lógicos que pueden ser BCCH, DCCH, PCCH, CCCH o DTCH.
 - b) Entidad TM RLC receptora, que recibe TMD PDU de la capa inferior a través del canal lógico configurado, y entrega RLC SDU a los niveles superiores.
2. UM (Unacknowledged Mode), que tendrá una entidad transmisora y otra receptora para cada servicio, y que utilizará un único canal lógico.
 - a) Entidad UM RLC transmisora, recibe RLC SDU de los niveles superiores, que pueden requerir o no de segmentación UMD PDU de tamaño apropiado. Si está configurado el cifrado, se cifra la UMD PDU (excepto la cabecera), antes de someterla a la capa inferior a través de los canales lógicos que pueden ser CCCH, DCCH, CTCH o DTCH.
 - b) Entidad UM RLC receptora, que recibe UMD PDU de la capa inferior a través del canal lógico configurado. Antes de entregar RLC SDU a los niveles superiores realiza el descifrado (si aplica), elimina la cabecera RLC y reensambla los RLC SDU si en la entidad UM RLC transmisora se ha aplicado segmentación y/o concatenación.
3. AM (Acknowledged Mode), que tendrá una única entidad transmisora y receptora combinada, y que puede estar configurada para utilizar uno o dos canales lógicos para enviar o recibir PDU de datos y control.
 - a) Parte transmisora, que recibe RLC SDU de la capas superior, las cuales se segmentan y/o concatenan en AMD PDU de longitud fija. El tamaño de los AMD PDU es semi-estático, se configura a través de capas superiores y sólo puede ser cambiado a través del reestablecimiento de la entidad AM RLS por parte de las capas superiores. Tras la operación de segmentación y/o concatenación, los AMD PDU se sitúan en el «buffer» de retransmisión. Allí bien se eliminan o se retransmiten según el reconocimiento positivo o negativo que se reciba de la unidad AM RLC pareada. Se incorpora la cabecera y si es el caso el cifrado, que no se aplica a la cabecera AMD PDU. La parte transmisora de la entidad AM RLC somete los AMD PDU a las capas bajas a través del canal lógico DCCH o DTCH.
 - b) Parte receptora, que recibe los PDU de datos y de control desde las capas bajas a través de los canales lógicos configurados, y se enrutan a la unidad de descifrado si aplica, antes de pasar al «buffer» de recepción. Los AMD PDU se mantienen en el buffer hasta la completa recepción del RLC SDU. El receptor reconoce la recepción exitosa o solicita la retransmisión a la unidad AM RLC pareada a través de la parte transmisora, enviando uno o más STATUS PDU. Cuando se consigue recibir por completo el RLC SDU, los AMD PDU asociados se reensamblan si se entregan a las capas superiores.

La interacción entre el subnivel RLC y los otros niveles se describe a través de primitivas, que representan el intercambio lógico de informaciones de control entre niveles.

E.7.3. MAC

La especificación del 3GPP correspondiente a este protocolo describe la arquitectura MAC y sus diferentes entidades desde el punto de vista funcional, pero no especifica ni restringe las implantaciones que se realicen del mismo. Las funciones que lleva a cabo el nivel MAC son las siguientes:

- ⇒ Modelado de canales lógicos sobre los canales de transporte.
- ⇒ Selección del TF apropiado para cada canal de transporte en cada instante.
- ⇒ Gestión de prioridades entre diferentes flujos de un mismo UE.
- ⇒ Gestión de prioridades entre diferentes UE a través de «scheduling» dinámico.
- ⇒ Identificación de UE sobre los canales de transporte comunes.
- ⇒ Medidas del volumen de tráfico.
- ⇒ «Transport Channel Type Switching».
- ⇒ Cifrado.
- ⇒ Control de las transmisiones sobre RACH y CPCH.

Los servicios que proporciona el MAC a las capas superiores son fundamentalmente los siguientes:

- ⇒ Transferencia de información en modo no reconocido entre entidades MAC pareadas, en forma de MAC-SDU.
- ⇒ Reasignación de recursos radio y cambio de los parámetros MAC, siguiendo las indicaciones que le lleguen desde el nivel RRC.
- ⇒ Reporte de medidas locales hacia el nivel RRC.

La interacción del nivel MAC con los otros niveles, en particular L1, RLC y RRC se lleva a cabo en términos de primitivas que representan el intercambio lógico de información y control entre los diferentes niveles.

E.7.4. PDCP

Los protocolos de la familia IP no se caracterizan especialmente por su eficiencia. La cabecera IP por sí sola (sin considerar posibles extensiones) consta de 20 bytes. Por encima de IP lo habitual es usar TCP o UDP, lo que supone añadir 20 u 8 octetos más, respectivamente. La sobrecarga debida a estas cabeceras puede llegar a ser un problema en escenarios donde el ancho de banda es un recurso escaso. Este es precisamente el caso de las redes celulares donde el espectro radioeléctrico disponible suele estar limitado. En este tipo de escenarios habría que añadir las sobrecargas introducidas por los protocolos radio (RLC y MAC), con lo que fácilmente se concluye lo ineficiente que puede ser el soporte de servicios IP.

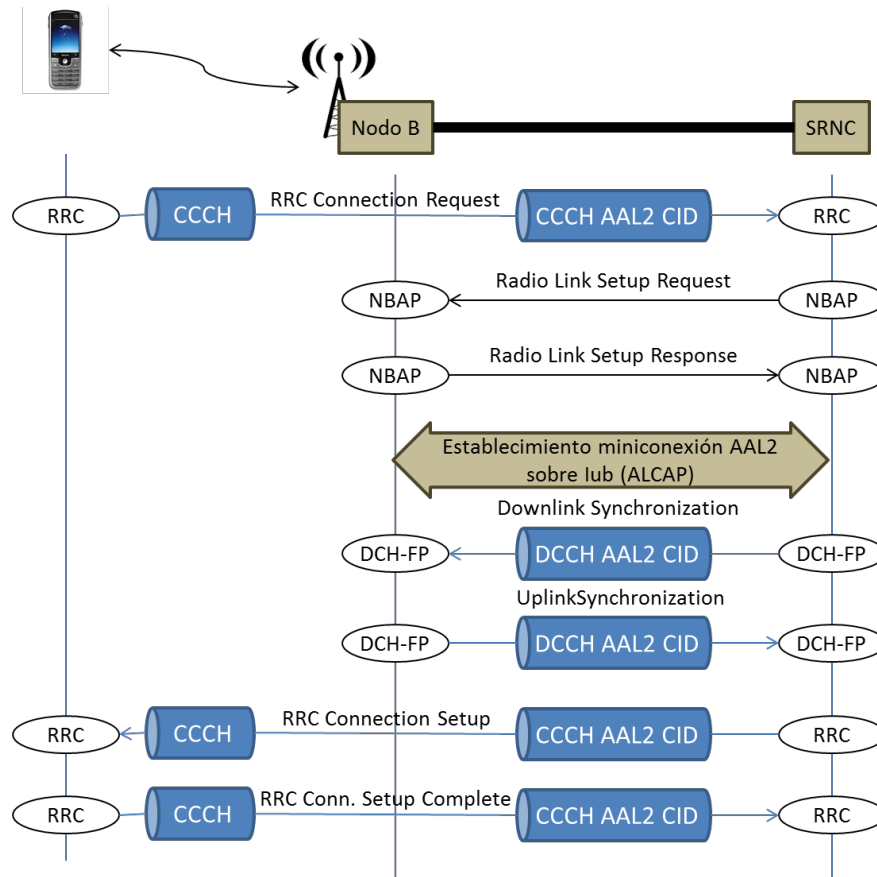


Figura E.17: Procedimiento de establecimiento de conexión RRC

mensaje RRC «Paging» al Nodo B donde se encuentra el usuario, sobre la miniconexión AAL2 de la interfaz Iub correspondiente al canal lógico de control dedicado (DCCH) del móvil.

E.8.2. Establecimiento de conexión RRC

A diferencia de lo que sucede en telefonía fija, en una red móvil los terminales no disponen de un medio de transmisión permanente (por ejemplo, el bucle telefónico) para intercambiar señalización con la red. Por este motivo, cada vez que el móvil desea entablar un diálogo con la red (para registrarse en la red, responder a un aviso, iniciar una llamada, actualizar su posición, etc), lo primero que ha de hacer es solicitar un canal de señalización. En UMTS, esta tarea se lleva a cabo mediante el procedimiento de establecimiento de conexión RRC, representado en la figura E.17.

El procedimiento se inicia mediante el envío por parte del terminal del mensaje RRC «Connection Request» sobre un canal lógico de señalización común (CCCH). El mensaje llega hasta SRNC vía la miniconexión AAL2 asociada a dicho canal radio, donde se procede a cursar la petición, ordenando vía NBAP la activación de un nuevo canal lógico dedicado (DCCH) para el móvil, así como el establecimiento de la correspondiente miniconexión AAL2 sobre Iub. Seguidamente, el RNC notifica al móvil, vía el canal de señalización común, cuál ha sido el canal que le ha sido asignado. Finalmente, el procedimiento termina mediante la confirmación que envía el móvil al RNC sobre el canal dedicado que el acaba de ser asignado.

Desde el punto de vista del acceso radio, el UE debe haber realizado una serie de tareas antes

de estar en disposición de iniciar el procedimiento de establecimiento de conexión RRC. y que se revisan a continuación, con el fin de proporcionar una visión más integradora de las diferentes entidades de la red, protocolos y procedimientos de nivel físico.

En primer lugar el UE debe estar sincronizado con la célula, esto es, haber detectado el canal de sincronismo SCH. El canal SCH permite, además del sincronismo temporal, el conocimiento del código de «scrambling» primario de la célula.

La información de «broadcast» (canal lógico BCCH, canal de transporte BCH) físicamente se transmite sobre el canal P-CCPCH, caracterizado por un código de canalización preestablecido (y por tanto conocido a priori por el UE) y el código de «scrambling» primario. Así, tras el proceso de sincronismo el UE está en disposición de leer la información de «broadcast».

Entre la información de «broadcast» se encuentra la configuración del canal de acceso en el enlace ascendente (RACH), indicándose por ejemplo el código de «scrambling» asociado al RACH. Con el código de «scrambling» del RACH y tras realizar la medida del nivel recibido del canal piloto CPICH (que al utilizar un código de canalización preestablecido y el código de «scrambling» primario, ya conocido por el UE, se está en disposición de detectar), se determina la potencia inicial a transmitir en el procedimiento de acceso.

Cuando se inicia el proceso de RRC «Connection Establishment», bien como respuesta a un «Paging» o bien por iniciativa de las capas superiores del UE, se lleva a cabo el proceso de acceso sobre el RACH. Cuando al UE se le reconoce, a través del canal AICH, el preámbulo de acceso, pasa a la transmisión de la parte del mensaje RACH (duración de 1 ó 2 tramas radio), transportando en particular el mensaje RRC «Connection Request». Entre los IE (Information Element) del mensaje se incluye el «Establishment cause» (por ejemplo «Originating Conversational Call», «Terminating Conversational Call», «Emergency Call», «Registration», «Detach», «Originating High Priority Signalling» o «Terminating High Priority Signalling».

Con el mensaje RRC «Connection Establishment», la SRNC está en disposición de decidir la asignación de recursos a esta transacción, así como decidir si se asignan recursos dedicados o compartidos.

En caso de asignarse recursos dedicados, se pasa a establecer el enlace entre la RNC y el Nodo B, mediante el mensaje «Radio Link Setup Request», que es el procedimiento empleado para establecer los recursos necesarios para un nuevo contexto de comunicación en el Nodo B. Con la recepción del mensaje el Nodo B reservará los recursos necesarios y configurará el nuevo «Radio Link(s)» según los parámetros especificados en el mensaje. Algunos de los IE que incluye son «UL DPCCH information», con información referente al canal de enlace ascendente (especificando por ejemplo el UL «scrambling code», código de «scrambling» que utilizará el UE, TFCS, posibles formatos de transporte que utilizará el UE, «UL DPCCCH Slot Format», «Power Offset Information», indicando la diferencia de potencia a aplicar para los bits de TFCL, TPC y piloto, y «FDD TPC DL Step Size» indicando el paso de incremento/decremento de potencia en dB asociado a un comando de control de potencia) así como «RL Information», información propia del enlace radio (RL ID, identificador del enlace, «Maximum DL Power», para indicar la potencia máxima asociada a la conexión).

Cuando la comunicación sobre la interfaz Iub está a punto, el mensaje utilizado por la red para notificar al UE la aceptación del establecimiento de la conexión RRC así como para indicar los parámetros seleccionados es el RRC «Connection Setup». Este mensaje incluye, además de los IE propios para la asignación de recursos radio, la asignación de un identificador a la transacción RRC y de una nueva RNTI (Radio Network Temporary Identity). El canal utilizado en este caso es el

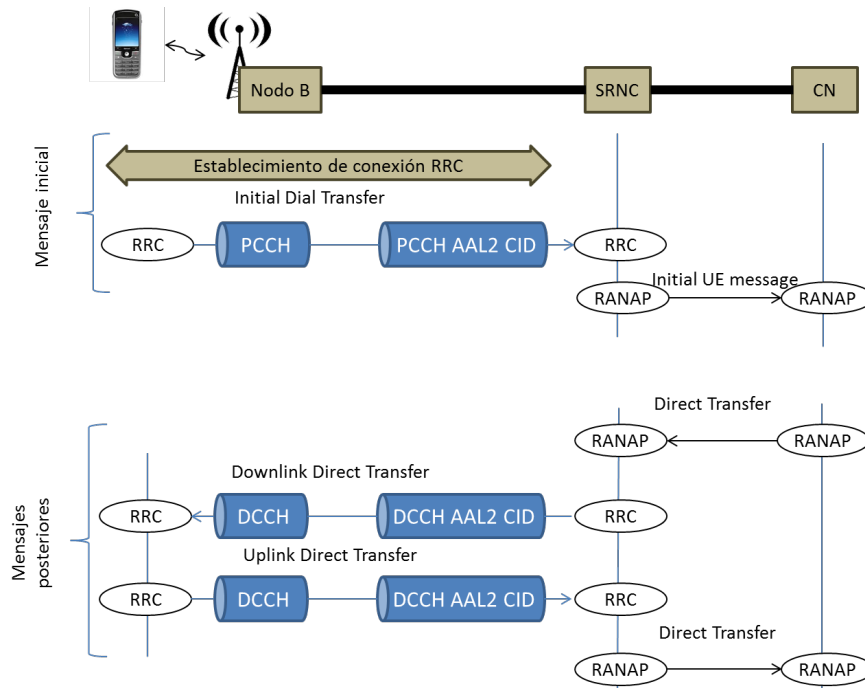


Figura E.18: Intercambio de señalización móvil-red sobre la UTRAN

FACH (canal común del enlace descendente).

Cuando el UE recibe el mensaje de RRC «Connection Setup» el nivel RRC configura las capas L1 y L2 de acuerdo a los parámetros indicados y establece localmente el canal lógico DCCH. Cuando los niveles L1 y L2 están correctamente configurados y establecidos, se envía el mensaje RRC «Connection Setup Complete» a la red para finalizar el procedimiento.

Nótese que a nivel de canal de transporte el canal lógico DCCH pueden materializarse a nivel de canal de transporte bien como DCH o bien como RACH/FACH.

E.8.3. Señalización móvil-red

En la figura E.18 se ilustra el intercambio de señalización móvil-red (estrato no ligado al acceso), el cual requiere la apertura previa de una sesión RRC (ver apartado anterior). La parte superior de la figura corresponde al envío del mensaje inicial por parte del móvil, que sirve para iniciar la conexión de señalización con el nodo de acceso al núcleo de red (MSC o SGSN) sobre la interfaz Iu (Iu-CS o Iu-PS). Dicho mensaje viaja sobre el canal de control dedicado (DCCH) y la miniconexión AAL2 correspondiente hasta alcanzar al SRNC. Este lo reenvía dentro del mensaje RANAP (Initial UE Message) hasta la MSC o el SGSN.

Los mensajes de señalización móvil-red subsiguientes, los cuales dependerán del motivo concreto por el que el móvil desea conectarse de la red, se llevan a cabo de acuerdo con las secuencias de mensajes representadas en la parte inferior de la figura. En la interfaz Iu, la señalización se intercambia por transferencia directa mediante el mensaje RANAP «Direct Transfer». En las interfaces Iub y Uu, se emplean los mensajes RRC «UL Direct Transfer» o «DL Direct Transfer» (según si el mensaje es de móvil a la red o al revés), transportados sobre la miniconexión AAL2 y el canal radio dedicado, respectivamente.

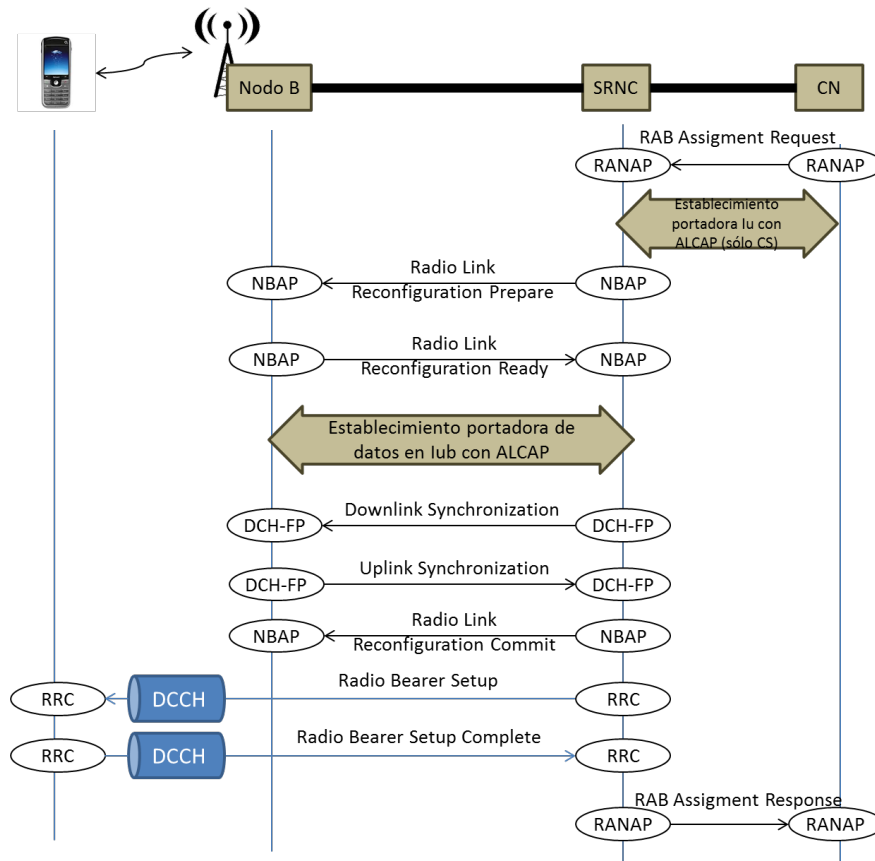


Figura E.19: Procedimiento de establecimiento de RAB

E.8.4. Establecimiento de portadora de acceso radio (RAB)

Como consecuencia del intercambio de señalización móvil-red, puede llegar un momento en el que la MSC o el SGSN decida establecer una portadora de acceso radio (RAB) para el intercambio de tráfico de usuario entre el móvil y el núcleo de red (por ejemplo, para una llamada de voz o una sesión de datos). Como se indica en la figura E.19, el procedimiento se inicia con el envío del mensaje RANAP «RAB Assignment Request» hacia el SRNC, el cual desencadena la asignación de recursos sobre los tres tramos que componen un RAB: la interfaz radio (a nivel lógico un canal de datos dedicado), la interfaz Iub (una miniconexión AAL2 para el transporte de las tramas MAC de dicho canal entre el Nodo B y el SRNC) y la interfaz Iu (un túnel GTP sobre una conexión AAL5 si se trata del dominio PS, o una miniconexión AAL2 si se trata del dominio CS).

En la figura E.20 se detalla el intercambio de mensajes de señalización AAL2 (o ALCAP, según la terminología del 3GPP) para el establecimiento de portadoras de datos sobre las interfaces Iub e Iu-CS. Nótese cómo es siempre el RNC el que envía el mensaje de establecimiento.

Nótese que el mensaje RAB «Assignment Request» lleva asociados los requisitos en cuanto a los distintos atributos de la RAB («Traffic Class», «RAB Asymmetry Indicators», «Maximum Bit Rate», «Guaranteed Bit Rate», «Delivery Order», «Maximum SDU Size», «SDU parameters», «Transfer Delay», «Traffic Handling Priority», «Allocation/Retention Priority», «Source Statistics Descriptor» y «Relocation Requirement»). Tras la llegada de este mensaje, es la función de control de admisión a UTRAN la que determina si hay suficiente capacidad en la interfaz radio para acomodar a la RAB solicitada y, en caso afirmativo, se prosigue el procedimiento.

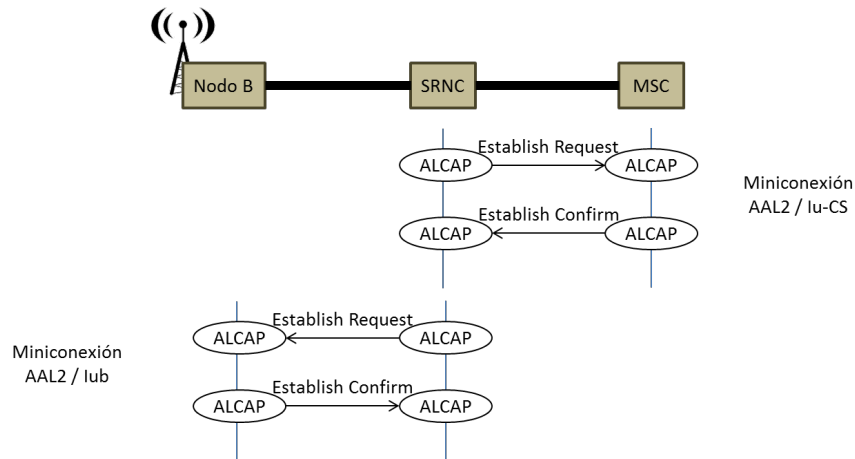


Figura E.20: Establecimiento de portadoras de datos AAL2

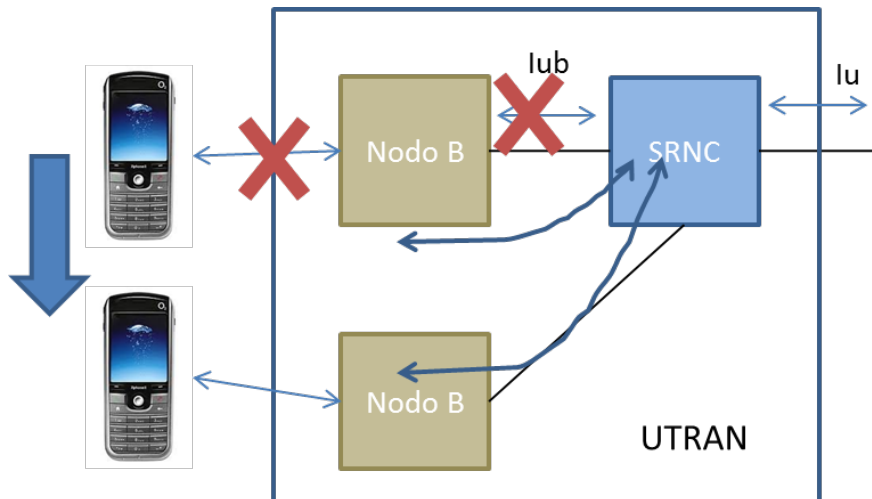


Figura E.21: Traspaso suave entre Nodos B controlados por el mismo RNC

E.8.5. Traspasos suaves («Soft Handover»)

Dentro de los procedimientos de UTRAN se incluyen los relativos a los traspasos, en sus distintas modalidades: traspasos «soft» y «hard», con o sin cambio de RNC, con o sin cambio de MSC o SGSN, traspasos entre células 2G y 3G, etc. De entre todos ellos, especial atención merecen los traspasos suaves («soft handover»), no sólo por su novedad (este tipo de traspasos no existe en GSM), sino por las implicaciones que tienen en cuanto al consumo de recursos en la red de acceso radio. En los apartados siguientes se ilustran algunos ejemplos de procedimientos de traspaso suave.

E.8.5.1. Traspaso suave sobre un mismo RNC

En la figura E.21 se ilustra un traspaso suave entre dos células pertenecientes a un mismo RNC. Al entrar en la región fronteriza entre células, y con objeto de minimizar la potencia requerida (y por tanto, la interferencia), el móvil se comunica simultáneamente con los dos Nodos B. Ello implica habilitar un segundo trayecto de comunicación hasta el SRNC: un canal radio en la nueva célula y la correspondiente miniconexión AAL2 sobre la interfaz Iub que lo conecta al SRNC.

Mientras el móvil permanece en la región de traspaso, se produce, en definitiva, un doble consumo

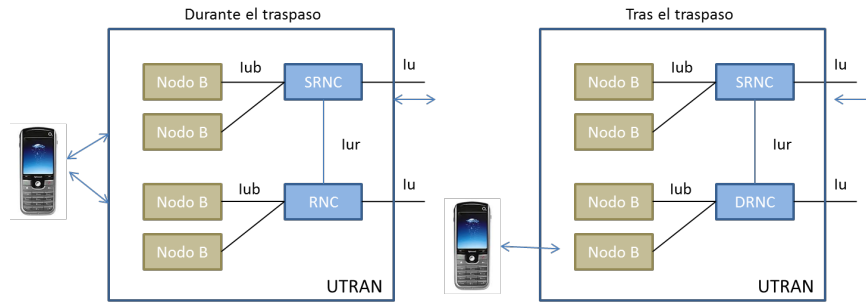


Figura E.22: Traspaso suave entre Nodos B controlados por distintos RNC

de recursos en UTRAN. Por motivos de eficiencia es necesario liberar los trayectos extra tan pronto el móvil abandona la zona de traspaso (por ejemplo, que el móvil retroceda hacia la célula de partida o se adentre en la nueva).

El procedimiento de traspaso suave entre dos células pertenecientes a un mismo RNC se realiza de la siguiente manera: inicialmente, el móvil se encuentra en la célula cubierta por el Nodo B número 1 (Nodo B-1). Al moverse, entra en la región fronteriza con la célula cubierta por el Nodo B número 2 (Nodo B-2), dando lugar al establecimiento de un segundo trayecto de comunicación (un nuevo canal radio en el Nodo B-2 más su correspondiente miniconexión AAL2 sobre Iub 2). Si el móvil se adentra en la nueva célula, se procede a liberar los recursos asociados al trayecto de comunicación original vía el Nodo B-1 (canal radio y miniconexión AAL2 sobre Iub-1).

E.8.5.2. Traspaso suave entre RNC

Otro ejemplo de traspaso suave es el que se da entre Nodos B controlados por distintos RNC. El trayecto secundario en este caso requiere, además del canal radio y la miniconexión AAL2 sobre la interfaz Iub del nuevo Nodo B, una miniconexión AAL2 entre el DRNC y el SRNC, esto es, sobre la interfaz Iur. La situación se muestra en la parte izquierda de la figura E.22.

Como en el ejemplo del apartado anterior, al abandonar la región de traspaso, el móvil vuelve a quedarse con un único trayecto de comunicación. Si ello se debe a que el móvil se adentra en la nueva célula, se llega a una situación como la representada a la derecha de la figura E.22, claramente ineficiente. Dicha ineficiencia se puede corregir mediante el procedimiento de reubicación de SRNS, descrito en el siguiente apartado.

El traspaso suave entre Nodos B pertenecientes a distintos RNC se realiza de la siguiente manera. Se crea el segundo trayecto de comunicación vía el DRNC (un nuevo canal radio, una miniconexión AAL2 sobre Iub2, y otra sobre Iur). Seguidamente, al adentrarse el móvil en la nueva célula, se procede a liberar el trayecto de comunicación a través del Nodo B inicial. El trayecto de comunicación resultante, a través de dos RNC, no es el más eficiente, siendo susceptible de ser optimizado mediante el procedimiento de reubicación de SRNS.

E.8.6. Reubicación de SRNS

El resultado del procedimiento de traspaso entre RNC puede dar lugar a trayectos de comunicación en UTRAN no óptimos, como el representado en la figura E.22. El procedimiento de reubicación de SRNS (SRNS «relocation») tiene por objeto precisamente la optimización de este tipo de trayectos, permitiendo que un DRNC pueda pasar a ser el nuevo SRNC para un cierto

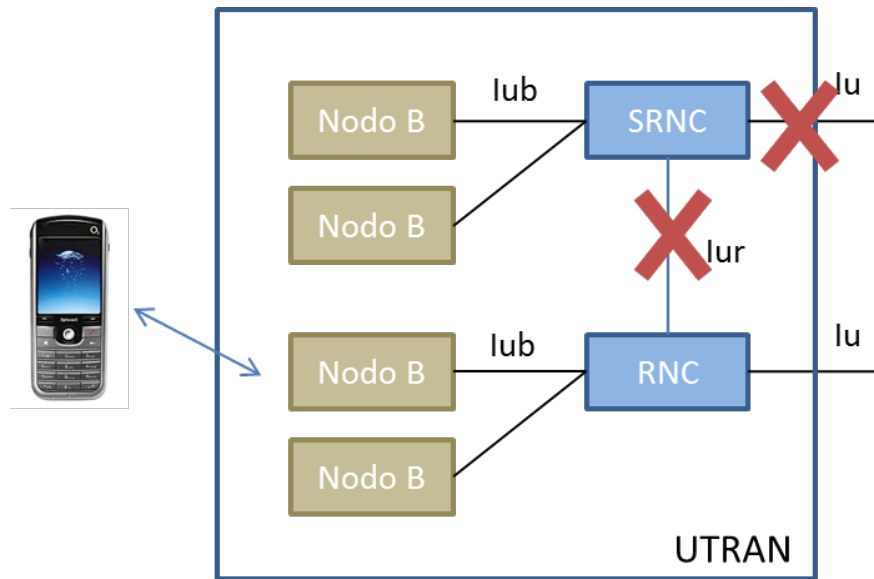


Figura E.23: Reubicación de SRNS

móvil, esto es, que sea su interfaz Iu la que soporte el intercambio de tráfico con el núcleo de red.

En la figura E.23 se ilustra el resultado del procedimiento de reubicación de SRNS. La reconfiguración del trayecto sobre la UTRAN, se evita la utilización de dos RNC y el consumo de recursos extra sobre la interfaz Iur.

El procedimiento de reubicación de SRNS comienza con el envío por parte del SRNC original del mensaje RANAP «Relocation Required» hacia el SMC (y/o SGSN, según las comunicaciones que el móvil tenga en curso). Suponiendo que el núcleo de red accede a la petición, la MSC (y/o SGSN) entabla contacto con el futuro SRNC (hasta el momento DRNC), solicitando la reserva de recursos sobre la nueva interfaz Iu. Una vez efectuada esta tarea, el MSC ordena al SRNC original que contacte con el DRNC (vía RNSAP) para completar el traspaso de papeles. El DRNC pasa a ser el nuevo SRNC, confirmándole al núcleo de red que la reubicación se ha completado. Finalmente, el núcleo de red ordena la liberación de recursos sobre la interfaz Iu original, que irá seguida también de la liberación de recursos en la interfaz Iur.

En el ejemplo se ha supuesto que el móvil está conectado solamente al dominio CS y que los dos RNC dependen de la misma MSC. En el informe TR 25.931 se proporcionan ejemplos más complejos.

E.9. Protocolos E-UTRAN

Se ha detallado ya en capítulos anteriores la arquitectura de la nueva tecnología LTE (*UMTS Long Term Evolution*). En los capítulos 2 y 3 se ha hecho referencia a la arquitectura de la red, sus distintas entidades y las interfaces que completan este nuevo esquema. En este apartado se pretende complementar dicha información explicando someramente los distintos protocolos que existen en cada una de las interfaces.

E.9.1. 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 E.24. 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, procesado 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. 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.

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:

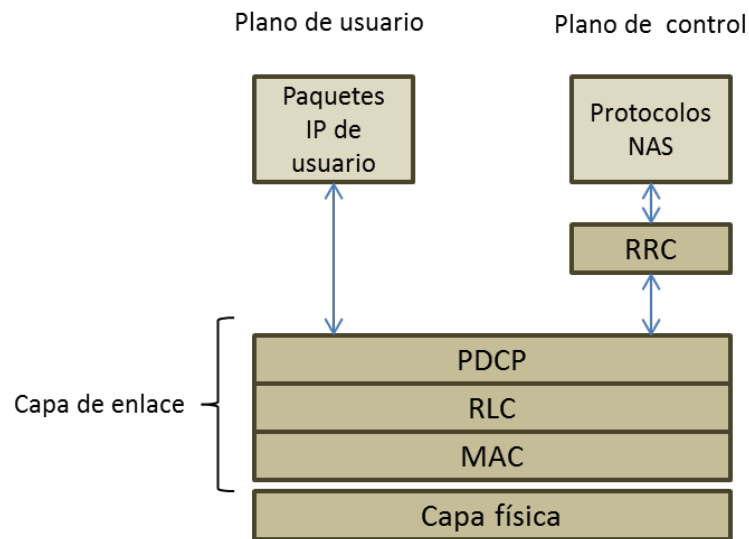


Figura E.24: Protocolos de la interfaz radio de E-UTRAN

- ⇒ *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.

E.9.2. 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 E.25 se ilustra la arquitec-

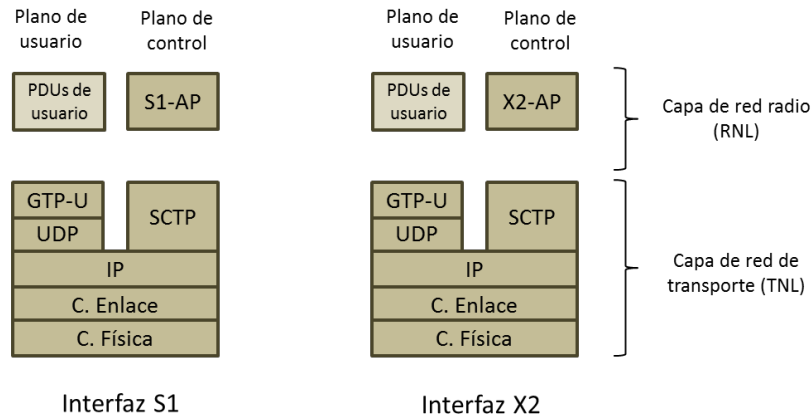


Figura E.25: Protocolos en las interfaces S1 y X2

tura de protocolos de las interfaces S1 y X2. Los documentos de base del 3GPP que especifican la estructura de las interfaces S1 y X2 son, respectivamente, 3GPP TS 36.410 y TS 36.420.

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 Iu-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—). 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 (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. 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 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

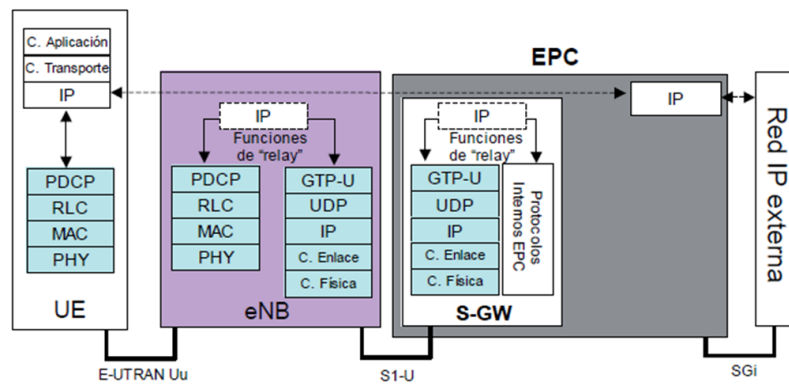


Figura 2.11 Protocolos del plano de usuario en E-UTRAN

Figura E.26: Protocolos del plano de usuario en E-UTRAN

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 se denomina X2-AP (*X2 Application Part*) y se especifica en 3GPP TS 36.423.

E.9.3. Plano de usuario entre UE y EPC

Atendiendo a la descripción realizada en los anteriores apartados, en la Figura E.26 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, RTP, 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.

E.9.4. Plano de control entre UE y EPC

En la Figura E.27 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 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.

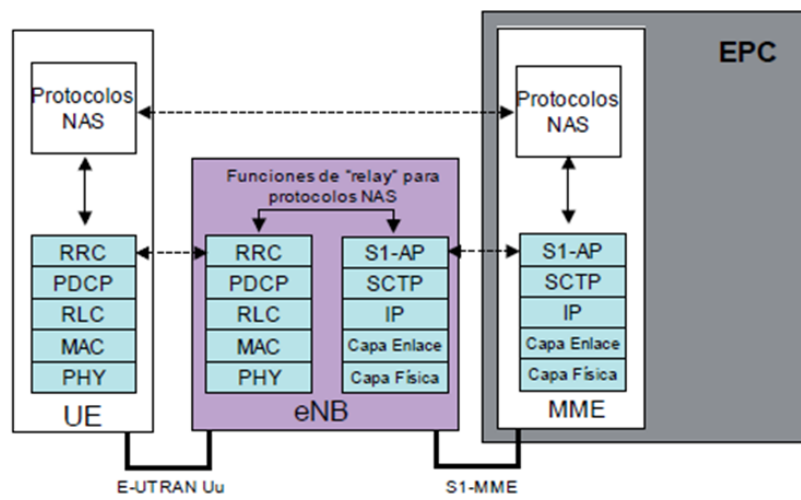


Figura 2.12 Protocolos del plano de control en E-UTRAN

Figura E.27: Protocolos del plano de control en E-UTRAN

Apéndice F

Notas sobre equipos Ericsson

F.1. KPIs Reales

A modo de ejemplo se presentan una serie de KPIs reales de la red Ericsson de un nodo B durante un día normal y que se utilizarán para validar el modelo propuesto de Iub.

- ⇒ **RAB-Att-CS**, indica el número de peticiones de cualquier tipo de RAB de circuitos efectuadas en un periodo de tiempo dado. Figura F.1.
- ⇒ **RAB-Att-PS-R99**, indica el número de peticiones de cualquier tipo de RAB de paquetes R99 efectuadas, es decir petición de RAB 64kbps, 128kbps o 384kbps, en un periodo de tiempo dado. Figura F.2.
- ⇒ **RAB-Att-HSDPA**, indica el número de peticiones RAB de HSDPA efectuadas en un periodo de tiempo dado. Figura F.3.
- ⇒ **RAB-Att-HSUPA**, indica el número de peticiones RAB de HSUPA efectuadas en un periodo de tiempo dado. Figura F.4.
- ⇒ **RAB-Fail-CS**, indica el número de peticiones rechazadas o falladas de cualquier tipo de RAB de circuitos efectuadas en un periodo de tiempo dado. Figura F.5.
- ⇒ **RAB-Fail-PS-R99**, indica el número de peticiones rechazadas o falladas de cualquier tipo de RAB de paquetes R99 efectuadas, es decir petición de RAB 64kbps, 128kbps o 384kbps, en un periodo de tiempo dado. Figura F.6
- ⇒ **RAB-Fail-HSDPA**, indica el número de peticiones rechazadas o falladas RAB de HSDPA efectuadas en un periodo de tiempo dado. Figura F.7.
- ⇒ **RAB-Fail-HSUPA**, indica el número de peticiones rechazadas o falladas RAB de HSUPA efectuadas en un periodo de tiempo dado. Figura F.8.
- ⇒ **Num-USER-hsdpa**, indica el número de clientes HSDPA activos por unidad de tiempo. Figura F.9.
- ⇒ **Num-USER-hsupa**, indica el número de clientes HSUPA activos por unidad de tiempo. Figura F.10.

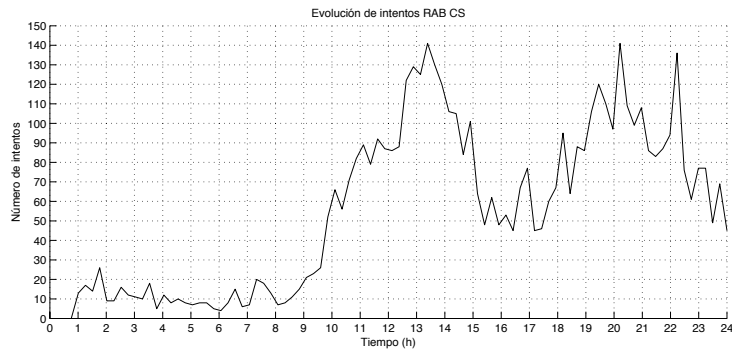


Figura F.1: Ejemplo del KPI 3G-ERI-RAB-Att-CS.

- ⇒ **TPUT-HS-P6**, indica la velocidad promedio HSDPA de todas las conexiones realizadas en el nodo.
- ⇒ **Minutes-Voice-q**, proporciona los minutos voz cursados durante un cuarto de hora.
- ⇒ **KB-HSDPA-DL-P6**, proporciona el tráfico en Kilobytes de paquetes HSDPA descargado por todos los usuarios. Figura F.13.
- ⇒ **KB-HSUPA**, proporciona el tráfico en Kilobytes de paquetes HSUPA subidos por todos los usuarios. Figura F.14.
- ⇒ **KB-PS-R99-DL**, proporciona el tráfico en Kilobytes de paquetes R99 descargado por todos los usuarios. Figura F.12.
- ⇒ **KB-PS-R99-UL**, proporciona el tráfico en Kilobytes de paquetes R99 subido por todos los usuarios. Figura F.12.
- ⇒ **SHO-factor**, proporciona el porcentaje de tráfico cursado en softhandover por unidad de tiempo. Figura F.15
- ⇒ **CSSR-CS**, proporciona un agregado de la accesibilidad total (complementario del bloqueo) de las conexiones sobre circuitos conmutados. Figura F.16.
- ⇒ **CSSR-PS**, proporciona un agregado de la accesibilidad total (complementario del bloqueo) de las conexiones sobre paquetes conmutados. Figura F.17.

El modelo de Iub simulará a partir de los KPIs de tráfico e intentos RAB, la accesibilidad del sistema cuando el factor limitando es el Iub. Se necesitará comparar los resultados que se obtengan con los KPIs reales CSSR-CS y CSSR-PS. El resultado final deberá ser similar ya que en estos dos KPIs se incluye el bloqueo por cualquier tipo de causa, bloqueo por Iub, por falta de CE, por falta de códigos o por un problema Hardware

F.2. Configuración de la red Ericsson

Una vez vistos los elementos que componen la red se detalla cómo se encuentran interconectados entre sí. Existen varias configuraciones dentro de una red Ericsson posibles.

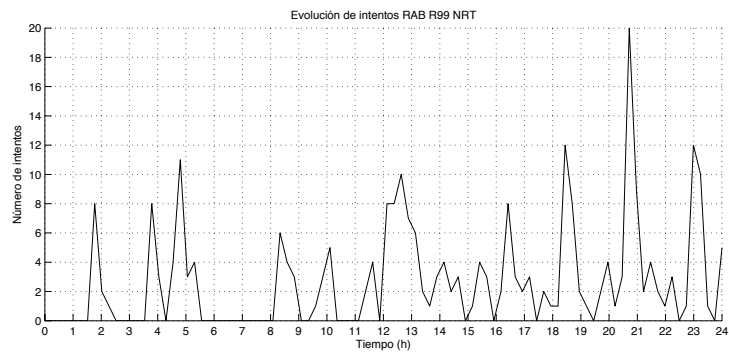


Figura F.2: Ejemplo del KPI 3G-ERI-RAB-Att-PS-R99

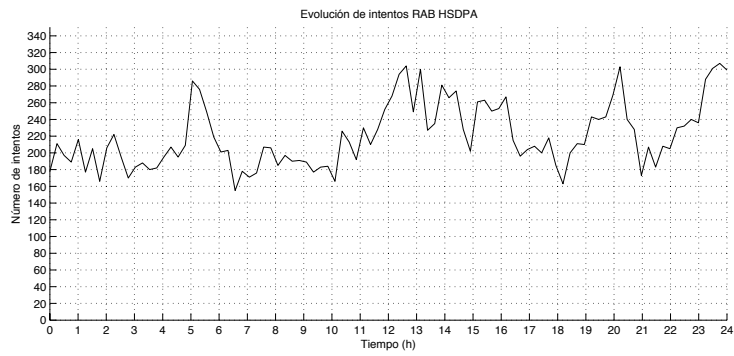


Figura F.3: Ejemplo del KPI 3G-ERI-RAB-Att-HSDPA.

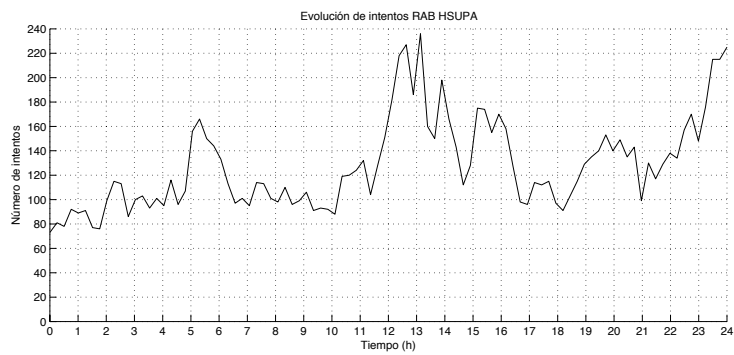


Figura F.4: Ejemplo del KPI 3G-ERI-RAB-Att-HSUPA.

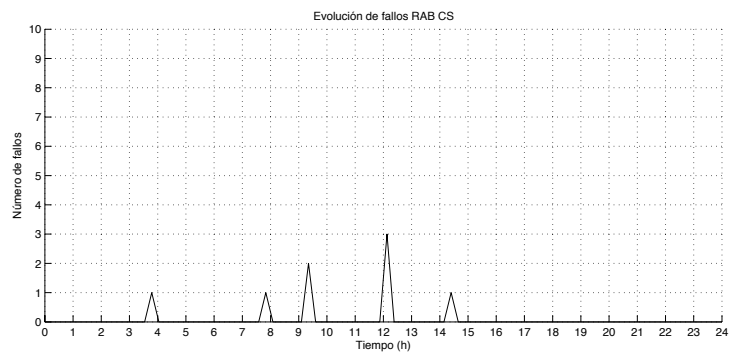


Figura F.5: Ejemplo del KPI 3G-ERI-RAB-Fail-CS.

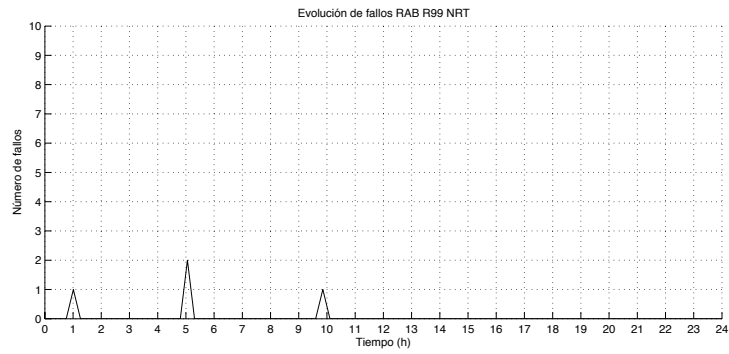


Figura F.6: Ejemplo del KPI 3G-ERI-RAB-Fail-PS-R99

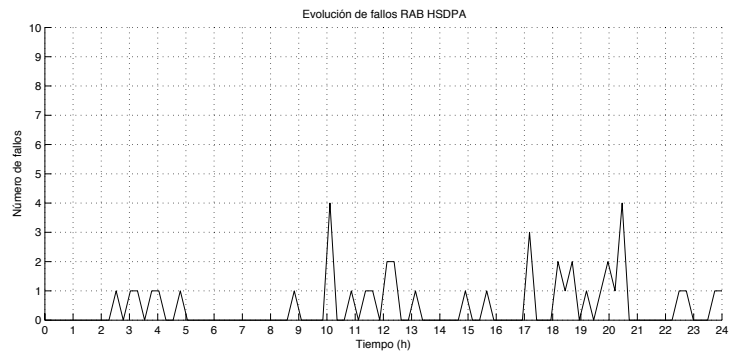


Figura F.7: Ejemplo del KPI 3G-ERI-RAB-Fail-HSDPA.

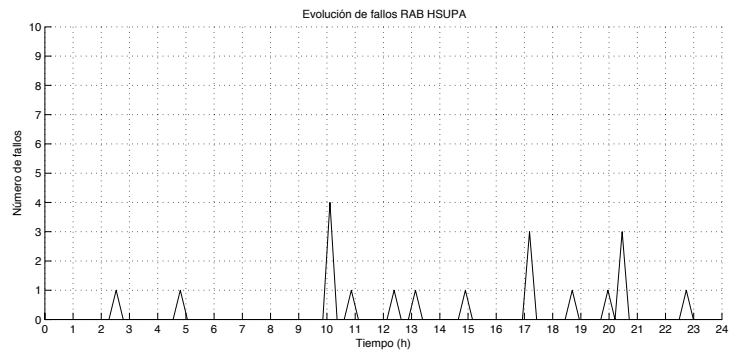


Figura F.8: Ejemplo del KPI 3G-ERI-RAB-Fail-HSUPA.

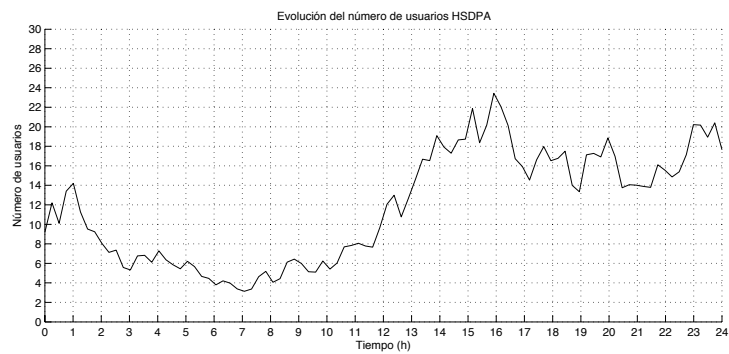


Figura F.9: Ejemplo del KPI 3G-ERI-Num-USER-hsdpa

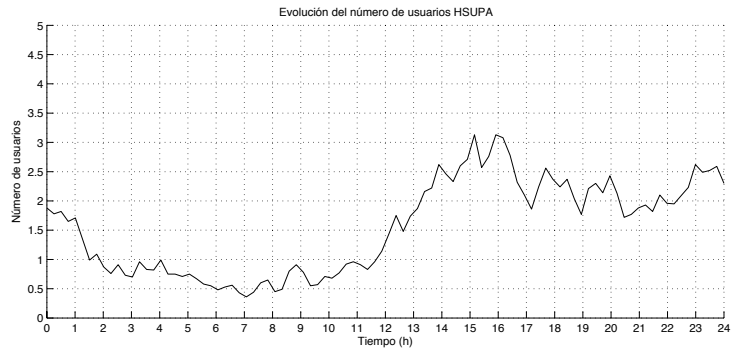


Figura F.10: Ejemplo del KPI 3G-ERI-Num-USER-hsupa.

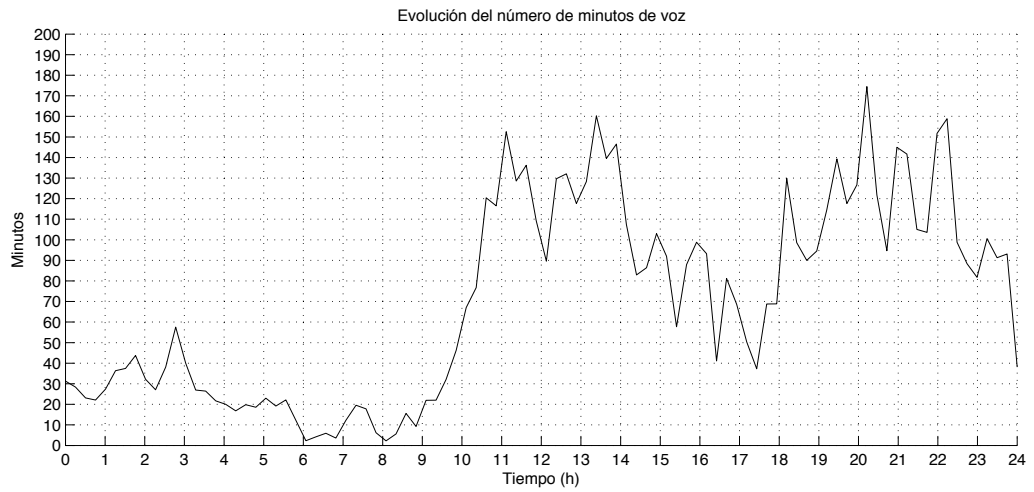


Figura F.11: Ejemplo del KPI 3G-ERI-Minutes-Voice-q.

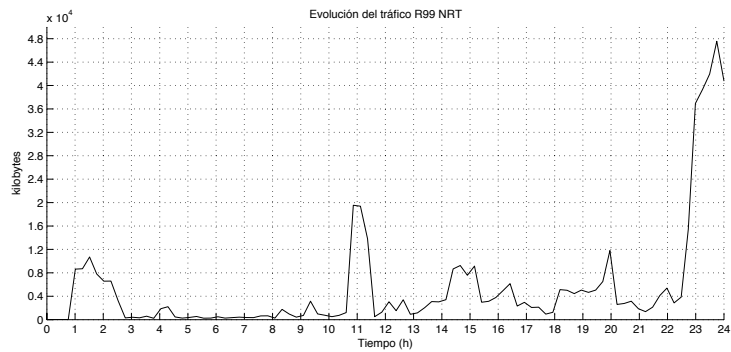


Figura F.12: Ejemplo del KPI 3G-ERI-KB-PS-R99.

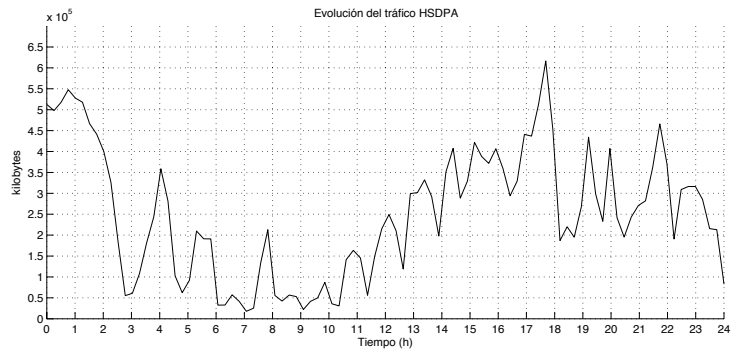


Figura F.13: Ejemplo del KPI 3G-ERI-KB-HSDPA-DL-P6

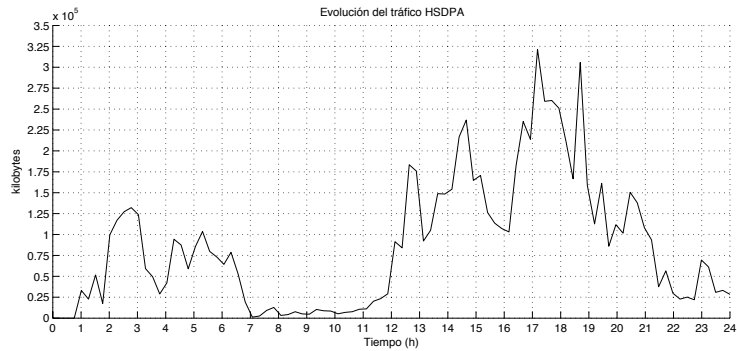


Figura F.14: Ejemplo del KPI 3G-ERI-KB-HSUPA.

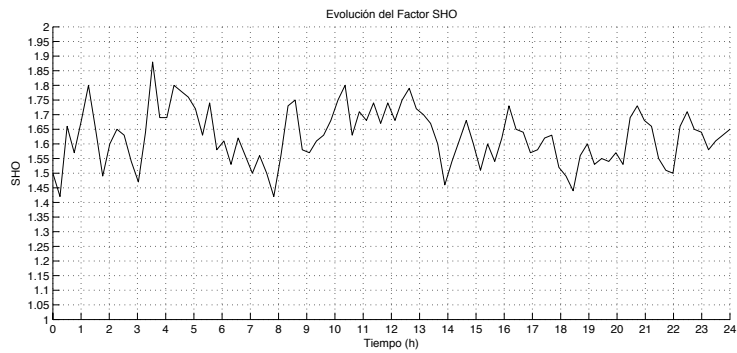


Figura F.15: Ejemplo del KPI 3G-ERI-SHO-factor

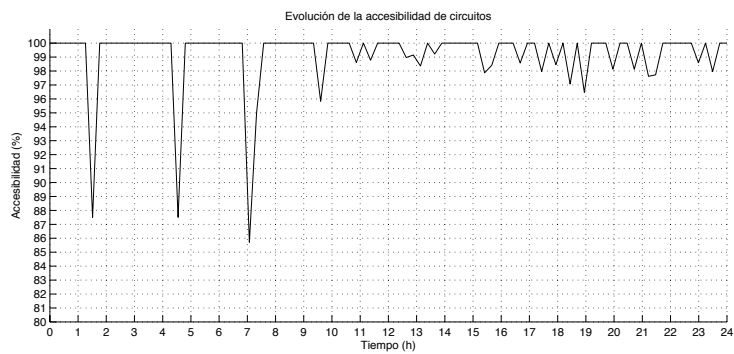


Figura F.16: Ejemplo del KPI 3G-ERI-CSSR-CS

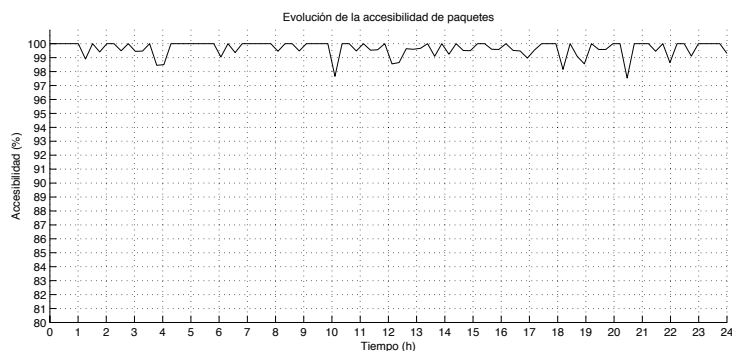


Figura F.17: Ejemplo del KPI 3G-ERI-CSSR-PS

	VPI	VCI	PCR (cps)	MCR (cps)	Clase	QoS
R99RT ¹	1	39	3000		Class 2	CBR
R99NRT ²	1	40			Class 4	UBR
HSPA	1	46			Class 4	UBR
NBAP-C	1	36	1000	160	Class 3	UBR+
NBAP-D	1	37	1000	160	Class 3	UBR+
Q.2630	1	38	1000	160	Class 3	UBR+
Synch	1	35	5	0	Class 1	CBR

Cuadro F.1: Descriptores de tráfico de los servicios sobre ATM

F.2.1. Nodos con transmisión ATM

Son nodos con transmisión totalmente basada en ATM y que, por tanto, todos los servicios viajan típicamente a través de un enlace PDH mediante un número N de E1s. Recuérdese que con la familia de RBS 3000 de Ericsson puede obtenerse hasta un máximo de 4 E1s si sólo está instalada una tarjeta CBU, a lo que podría ampliarse con otros 8 E1s con tarjetas ETMC opcionales.

Ericsson implementa también en ATM configuraciones denominadas *AAL2 switching*. Esta técnica permite multiplexar conexiones AAL2 procedentes de varios VCs AAL2 entrantes sobre uno o varios VCs AAL2 de salida. Los beneficios de emplear esta técnica son los siguientes:

- ⇒ El uso de *AAL2 switching* en nodos intermedios reduce el número de multiplexadores AAL2 en la RNC.
- ⇒ Emplear *AAL2 switching* en *hubs* reduce el ancho de banda requerido y, por tanto, supone un ahorro de líneas de transmisión.
- ⇒ Cuando se emplean hubs con soporte *AAL2 switching* se tiene la posibilidad de ofrecer redundancias a través de balanceos de carga cuando existen dos caminos físicos hacia la RNC.

Independientemente de si se utiliza transmisión ATM normal como mediante *AAL2 switching* los descriptores de tráfico definidos para la red serán los mismos, reflejados en la tabla F.1.

F.2.2. Nodos con transmisión *DualStack*

La configuración *DualStack* es la nomenclatura empleada por Ericsson para designar a aquellos nodos con transmisión híbrida, es decir, nodos en los que unos servicios se cursan a través de tecnología ATM y otros a través de tecnología IP.

Servicio	Tecnología
Voz	ATM
Videollamada	ATM
Datos R99 NRT	ATM
Datos HSDPA	IP
Datos HSUPA	ATM
Señalización NBAP-C	ATM
Señalización NBAP-C	ATM
Señalización Q.2630	ATM
Sincronización	ATM

Cuadro F.2: Configuración DualStack.

En la red que ocupa a esta Tesis los nodos *DualStack* presentan todos una configuración homogénea, tal y como se muestra en la tabla F.2.

F.2.3. Nodos Full-IP

Dentro de esta categoría se encuentran nodos cuyos servicios se transportan íntegramente sobre conexiones IP; la gestión de prioridades se realiza mediante el uso de la byte de prioridad de las cabeceras IP.

En la red UTRAN estudiada no existían nodos totalmente Full-IP, sino que empleaba la versión P6 de Ericsson. En ella, los nodos B requieren de un E1 para sincronización, el resto de servicios pueden viajar a través del camino IP.

F.3. Caracterización de recursos en equipos Ericsson

En el capítulo 5 se han presentado desde un punto de vista general los diferentes recursos que se han tenido en cuenta en el modelado a lo largo de la Tesis. La escasez o no de los mismos da como resultado la aparición de factores que limitan la accesibilidad del sistema total provocando la degradación del mismo. En particular, uno de los recursos considerados ha sido la capacidad de procesado en banda base del Nodo B, medido en unidades de Channel Elements. En dicho capítulo se comentó que la provisión de este recurso se realiza en la red mediante dos tipos de tarjetas, unas específicas para el enlace descendente (DL) y otras para el ascendente (UL). Para el caso particular de equipos Ericsson, se tiene:

⇒ Tarjeta HSTX para el enlace descendente.

⇒ Tarjeta RAXB para el enlace ascendente.

La gestión de los ResourceIds es algo diferente según el modelo concreto de tarjeta.

HSTX45

⇒ Se parte de 384 CE.

⇒ Si se activa HSUPA: -128 CE.

⇒ Si se activa segundo ResourceId: -128 CE.

⇒ Si se activa tercer ResourceId: -128 CE. Cada 16 usuarios HSDPA: -21 CE.

HSTX60

Se parte de 384 CE.

⇒ Por defecto se activa HSUPA.

⇒ Si se activa segundo ResourceId: -0 CE.

⇒ Si se activa tercer ResourceId: -128 CE.

⇒ Cada 16 usuarios HSDPA: -21 CE.

Un ejemplo del calculo de los CE reservados por los servicios HSPA podra ser el siguiente:

⇒ Se necesitan 2 ResourceIds para HSDPA.

⇒ Se desea activar el servicio en sentido ascendente (HSUPA).

⇒ 48 usuarios.

Los cálculos quedan de la siguiente manera:

⇒ HSTX45. $384\text{CE} - 128\text{CE}(\text{HSUPA}) - 128\text{CE}(\text{2o ResId}) - 48/16 \times 21\text{CE}(48 \text{ usuarios}) = 65\text{CE}$ disponibles

⇒ HSTX60. $384\text{CE} - 0\text{CE}(\text{2o ResId}) - 48/16 \times 21\text{CE}(48 \text{ usuarios}) = 321\text{CE}$ disponibles

F.4. Descripción de los casos de uso

F.4.1. Características generales de los casos de uso

Modelo de nodo: Ericsson 3000. El modelo de estación base es compartido para todos los casos de estudio analizados. No así la configuración del mismo. Entre los factores de configuración comunes a todos los casos de estudio analizados se encuentran:

⇒ Factor de potencia.

⇒ Recogida y procesamiento de KPI.

⇒ Modelos de tarjetas CE HSTX45, HSTX60, el número podrá variar.

⇒ Modelos de tarjetas CE RAX14, el número de ellas podrá variar en los distintos escenarios.

⇒ Posibilidad de configuración Dual Stack o Dual IP.

⇒ Enlaces ATM E1 (2Mbps).

⇒ Enlaces IP 23Mbps.

F.4.2. Particularidades de los casos de uso

F.4.2.1. Caso de uso A, Cuadro F.3

En este nodo se decide ampliar la capacidad de la interfaz Iub en el UL, un VCC.

Este nodo está configurado con 2 módulos RAX13, 1 módulo HSTX45 y otro HSTX60. Tiene dos portadoras diferenciadas y un enlace DL IP de 23 Mbps. En el primer escenario, configurado con un único VCC en el UL, se observan degradaciones de la accesibilidad. Como respuesta se amplía un VCC adicional. Los resultados de las dos situaciones serán monitorizados en las condiciones expuestas anteriormente.

Escenario/Recursos	CE UL	CE DL	SC	Iub UL	Iub DL
1	2xRAX 13	1xHSTX45 + 1xHSTX60	2 portadoras	1 VCC	IP 23 Mbps
2	2xRAX 13	1xHSTX45 + 1xHSTX60	2 portadoras	2 VCC	IP 23 Mbps

Cuadro F.3: Recursos para cada escenario

F.4.2.2. Caso de uso B Cuadro F.4

Este nodo presenta un comportamiento similar al anterior, y se decide ampliar la capacidad de la interfaz Iub en el UL, pasando de uno a dos VCC.

Escenario/Recursos	CE UL	CE DL	SC	Iub UL	Iub DL
3	2xRAX 13	1xHSTX45 + 1xHSTX60	2 portadoras	1 VCC (5 E1)	IP 23 Mbps
4	2xRAX 13	1xHSTX45 + 1xHSTX60	2 portadoras	2 VCC (5 E1)	IP 23 Mbps

Cuadro F.4: Recursos para cada escenario

F.4.2.3. Caso de uso C Cuadro F.5

Este nodo presenta congestión debida fundamentalmente a la escasez de recurso Channel Elements, por lo que se decide incorporar una nueva tarjeta RAX 14 en el UL.

Escenario/Recursos	CE UL	CE DL	SC	Iub UL	Iub DL
5	2xRAX 14	2xHSTX45	1 portadoras	1 VCC (1 E1)	IP 30 Mbps
6	3xRAX 14	2xHSTX45	1 portadoras	1 VCC (1 E1)	IP 30 Mbps

Cuadro F.5: Recursos para cada escenario

F.4.2.4. Caso de uso D Cuadro F.6

A diferencia de los tres casos anteriores, donde la configuración de los nodos es Dual-Stack, se presenta aquí un nodo «Full-IP», es decir, aquel en el que la transmisión en la interfaz Iub se implementa mediante IP tanto en el enlace ascendente como en el descendente. Analizando el comportamiento del nodo para el tráfico ofrecido, se aprecia congestión debida a la escasez del recurso Channel Elements, por lo que se decide añadir una nueva tarjeta RAX14.

Escenario/Recursos	CE UL	CE DL	SC	Iub UL	Iub DL
7	2xRAX 14	2xHSTX45	1 portadoras	IP 30 Mbps	IP 10 Mbps
8	3xRAX 14	2xHSTX45	1 portadoras	IP 30 Mbps	IP 10 Mbps

Cuadro F.6: Recursos para cada escenario

Apéndice G

Técnicas de aprendizaje e inteligencia

G.1. Introducción

Uno de los recursos con que cuentan los operadores de redes móviles para la optimización de sus recursos son las estadísticas reportadas por sus nodos. A lo largo de los capítulos previos se ha ido desarrollando una aproximación a un modelo teórico, incluyendo la calibración y la validación del mismo en diferentes escenarios. Para la calibración del modelo se ha utilizado un modelo realimentado, manteniendo bajo control la calidad del sistema y tratando de ajustar iterativamente en diferentes supuestos las variables del modelo.

El propósito de este capítulo es mostrar al lector diferentes técnicas de aprendizaje como recursos para la calibración de un modelo similar al que se ha presentado a lo largo de esta Tesis, poniendo de manifiesto en qué contexto puede aplicarse cada uno de ellos, cuáles son sus puntos clave y su conveniencia. El autor presenta asimismo los respectivos planteamientos teóricos de cara a una futura implementación en este modelo.

Concepto de Inteligencia Colectiva

La expresión «inteligencia colectiva» representa un concepto que ha sido utilizado durante décadas, y cuya popularidad ha experimentado un notable crecimiento con la llegada de las nuevas tecnologías de la comunicación. Así, bajo esta expresión subyace la combinación de varios factores, como comportamientos, preferencias e ideas de grupo que permiten exponer nuevos puntos de vista [97].

Dos de los escenarios más básicos en los que se aplican técnicas de inteligencia colectiva son las encuestas y los censos. La recolección de respuestas de un gran número de personas permite esbozar conclusiones sobre el grupo a nivel estadístico, algo que de manera individual ningún miembro podría conocer por sí solo. Así, la inteligencia colectiva representa el esbozo de nuevas conclusiones a partir de contribuyentes independientes.

Un ejemplo perfectamente conocido son los mercados financieros, donde el precio no está fijado por un individuo o por un esfuerzo coordinado, sino por el comportamiento de varias personas independientes todas ellas actuando de acuerdo con lo que consideran de su interés. A priori puede parecer intuitivo pero mercados en los cuales varios participantes alcanzan acuerdos basados en sus creencias acerca de cómo serían los precios futuros, son considerados más eficaces a la hora de predecirlos frente al trabajo de expertos que hagan proyecciones de forma independiente. Todo ello

es debido a que este tipo de mercados combinan el conocimiento, la experiencia y el punto de vista de miles de personas para crear un pronóstico mejor que el que pueda delegarse en la perspectiva de una sola persona.

Aprendizaje automático

El aprendizaje automático es un campo de la inteligencia artificial que se ocupa de la algorítmica necesaria para permitir el aprendizaje de un computador. En la mayoría de los casos, esto significa que se trata de un algoritmo al que se le entregan un conjunto de datos e infiere información acerca de las propiedades de los mismos; esa información permite hacer predicciones acerca de otros datos que podrían obtenerse en un futuro. Todo ello es posible puesto que casi todos los datos de carácter no aleatorio contienen patrones, y éstos permiten a la máquina hacer generalizaciones [97]. Para ello, se entrena un modelo con lo que determina que son las características más importantes de los datos.

Existen un gran número de algoritmos de aprendizaje automático diferentes, todos con diferentes puntos fuertes y adecuados a diferentes tipos de problemas. Algunos, como los árboles de decisión, son transparentes de modo que el usuario puede entender totalmente cuál es el razonamiento llevado a cabo por la máquina. Otros, como las redes neuronales [96], son cajas negras: producen una respuesta pero es muy difícil reproducir cuál es el razonamiento subyacente.

El aprendizaje automático tiene también una serie de puntos débiles. Los algoritmos varían en su capacidad de generalizar sobre conjuntos de patrones grandes, y un patrón que no se parezca a ninguno que haya visto el algoritmo previamente es bastante probable que sea malinterpretado [97]. Mientras que los humanos tienen como recurso un cierto grado de conocimiento cultural y experiencia, así como una reseñable habilidad para reconocer situaciones similares cuando toman decisiones acerca de información nueva, los métodos de aprendizaje automático solo pueden hacer generalizaciones basadas en los datos que ya han visto e incluso en ese caso lo hacen de una forma limitada.

G.2. Modelado mediante árboles de decisión

G.2.1. Concepto y estructura de un árbol de decisión

Un árbol de decisión es un diagrama que representa de forma secuencial condiciones y acciones; muestra qué condiciones se consideran en primer lugar, cuáles en segundo y así sucesivamente. Este método permite mostrar la relación que existe entre cada condición y el grupo de acciones permisibles asociado con ella. Un árbol de decisión sirve para modelar funciones discretas, en las que el objetivo es determinar el valor combinado de un conjunto de variables, y basándose en el valor de cada una de ellas, determinar la acción a ser tomada.

Los árboles de decisión son normalmente construidos a partir de la descripción de la narrativa de un problema. Éstos proveen una visión gráfica de la toma de decisión necesaria, especifican las variables que son evaluadas, qué acciones deben tomarse y el orden en la cual se efectuará la toma de decisión. Cada vez que se ejecuta un árbol de decisión, se seguirá siempre un único camino, dependiendo del valor actual de la variable evaluada.

La estructura de un árbol de decisión sigue el siguiente esquema:

⇒ Cada nodo del árbol es una pregunta.

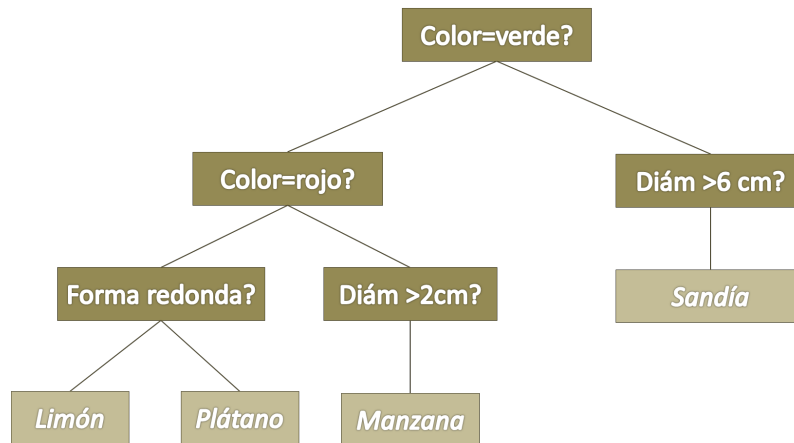


Figura G.1: Ejemplo de árbol de decisión

- ⇒ Cada nodo tiene tantos hijos como posibles respuestas haya para esa pregunta.
- ⇒ Las hojas del árbol corresponden a las distintas clases o categorías en las que las instancias son clasificadas por el concepto objetivo de acuerdo a las respuestas obtenidas desde la raíz hasta la hoja.
- ⇒ El espacio de hipótesis es entonces una disyunción de conjunciones sobre las características (es decir, permitiría representar el mayor conjunto de hipótesis sin sesgo inductivo).

Un ejemplo de árbol de decisión puede verse en la figura G.1.

G.2.2. Construcción de un árbol de decisión

En la mayoría de las ocasiones, los datos con los que se trabaja están organizados (o se pueden organizar) en un formato de base de datos, formado por registros y campos.

- ⇒ Un campo es una característica o atributo. Por ejemplo, en una base de datos de trabajadores sexo, edad, nacionalidad y departamento. Desde un punto de vista más abstracto, vectorial, correspondería a la definición del vector y sus dimensiones, tipología de datos, etc.
- ⇒ Un registro es un conjunto de datos particularizado, un vector. En el ejemplo, la definición de un individuo concreto a partir de 4 atributos.

Un árbol de decisión trata de establecer algún tipo de conclusión a partir de la respuesta secuencial a preguntas o condiciones impuestas sobre cada uno de los campos, que permiten así descartar cierto n° de registros y continuar el análisis con subconjuntos más pequeños.

A partir de la información de partida, se trata de establecer en qué orden deben ser interrogados los campos, o dicho de otra manera, cuál es en cada momento el mejor atributo para hacer esa clasificación.

La construcción del árbol puede realizarse conforme al siguiente procedimiento recursivo [97]:

1. Se parte de un árbol vacío.
2. Se consideran los atributos no utilizados.
3. Se elige al que mejor discrimina entre los ejemplos.
4. Para cada valor del atributo se crea un nuevo árbol utilizando el mismo método.

X=Verdadero	X=Falso
Y=Básico	Y=Ninguno
Y=Ninguno	Y=Ninguno
Y=Premium	Y=Básico
Y=Básico	Y=Premium
Y=Ninguno	Y=Básico

Cuadro G.1: Ejemplo de variable discriminatoria ineficaz

E>25	E<25
Y=Básico	Y=Ninguno
Y=Básico	Y=Ninguno
Y=Básico	Y=Ninguno
Y=Básico	Y=Premium
Y=Ninguno	Y=Premium

Cuadro G.2: Ejemplo de variable discriminatoria eficaz

G.2.3. Entrenamiento del árbol

Como ya se ha visto anteriormente, la elección de la mejor variable discriminatoria permite construir un árbol eficaz, ya que segmenta la información y permite obtener conjuntos más pequeños que posibilitan una toma de decisiones rápida y eficiente. En el diseño de los algoritmos, resulta de la misma importancia (si no más) la elaboración de un buen árbol antes de continuar con el proceso de implementación.

Para el entrenamiento del árbol se consideran todos los datos de entrada y de salida. Durante el entrenamiento del árbol se trabajará con los datos de entrada, se recorrerá el árbol y se comparará el resultado obtenido teórico con el real.

G.2.3.1. Método de elección de la mejor variable discriminatoria

Considerar en una de las etapas del proceso de diseño a una variable como la mejor, significa que los dos subconjuntos en que queda dividida la información de entrada dan como resultado variables de salida lo menos mezcladas posible [97]. La tabla G.1 muestra un ejemplo. Se tiene información sobre una serie de variables que determinan el tipo de suscripción que realizará un visitante a una web de servicios de pago. Se utiliza la variable «X» como variable discriminatoria y se divide la información inicial en 2 subconjuntos. Para cada uno de los registros se muestra la variable de salida «Y», que describe el tipo de suscripción. Como se puede observar, «X» no es en ese momento la mejor opción como variable discriminatoria, ya que las variables de salida no quedan lo suficientemente separadas (la mezcla de resultados es notoria). Sin embargo, en la tabla G.2 se ha utilizado «E» como variable discriminatoria. En este caso, los registros quedan divididos de una forma mucho más eficaz, ya que ha desaparecido la mezcla de resultados que se observaban en la tabla G.1, los resultados aparecen mucho más separados.

La «calidad» de esta división de la información, en base a una decisión acerca del valor de una variable se ha hecho de una forma cualitativa mediante un conteo del nº de veces que aparece repetido un determinado valor de la variable de salida. Hay varias formas de representar se una forma más cuantitativa cómo de mezclados están los registros dentro de un conjunto. Las dos métricas más conocidas son la «impureza de Gini» y la «entropía».

G.2.3.2. Impureza de Gini

El coeficiente, índice o relación de Gini es una medida de dispersión estadística desarrollada por el estadístico y sociólogo italiano Corrado Gini, publicado en 1912. Supóngase que se extrae un elemento aleatorio de un conjunto de datos y supóngase también que se etiqueta de forma aleatoria de acuerdo con la distribución de etiquetas de un subconjunto. La impureza de Gini es una medida que representa con qué frecuencia esa muestra será clasificada incorrectamente, y puede calcularse de acuerdo con la expresión G.1:

$$0 \leq GI \leq 1 - \sum_{i=1}^k p_i^2 \quad (\text{G.1})$$

donde k es el n.º de clases posibles, p_i el porcentaje de registros que hay en la clase i o su frecuencia y GI es el coeficiente de Gini.

Si todos los registros se encuentran perfectamente clasificados y no hay mezcla alguna, el coeficiente toma su valor mínimo (0). Por el contrario, a medida que su valor aumenta significa que los datos de las k clases se pueden encontrar con frecuencias cada vez más iguales dentro de ese conjunto.

G.2.3.3. Entropía

En teoría de la información, la entropía es una medida de cómo de desordenado se encuentra un conjunto o, dicho de otro modo, cómo de mezclado se encuentra. La expresión matemática que describe la entropía se representa en la ecuación G.2.

$$E = - \sum_{i=1}^k p_i \log_2(p_i) \quad (\text{G.2})$$

donde k es el n.º de clases posibles, p_i el porcentaje de registros que hay en la clase i o su frecuencia y E es la entropía. El valor de la misma se encuentra acotado entre dos límites, tal y como indica la expresión G.2.

$$0 \leq E \leq \log_2(k) \quad (\text{G.3})$$

Un valor de entropía de cero implica que el conjunto está ordenado. Cuanto más mezclados estén los grupos, mayor será el valor de la entropía. Por tanto, el objetivo a la hora de elegir cuál es la mejor variable discriminatoria es elegir aquella que proporcione el mínimo valor de entropía posible.

Si se establece una comparación entre las dos métricas (impureza de Gini y entropía) realizando diversos experimentos, puede comprobarse que la entropía alcanza su punto máximo de una forma mucho más lenta. Por tanto, tiende a penalizar las mezclas de una forma mucho más fuerte. Aunque cualquiera de las dos métricas serviría para el análisis que se pretende realizar, se utilizará de ahora en adelante la entropía por ser aquella que se utiliza más frecuentemente.

G.2.4. Construcción recursiva del árbol de decisión

Para comprobar cómo de buena es una variable concreta, el algoritmo debe calcular previamente la entropía del conjunto completo. Después, dividirá el conjunto original en varios subconjuntos de

acuerdo con el criterio de decisión seleccionado para esa variable, y se calculará la entropía para cada uno de ellos.

Se puede definir como «Ganancia de Información» la diferencia entre la entropía actual y la entropía media ponderada de los dos subconjuntos. De esta manera, para determinar cuál es la mejor variable discriminatoria para dividir el conjunto original se calculará previamente el factor de ganancia para cada una de las variables y se elegirá de entre ellas aquella que presente un valor de ganancia mayor.

Después de que la condición raíz del árbol ha sido determinada, ésta divide al árbol en al menos dos ramas. Las observaciones se distribuyen por tanto entre las que cumplen ese criterio y las que no lo cumplen. Para cada rama, el algoritmo determina si puede seguir siendo dividida o si por el contrario se ha alcanzado una conclusión sólida. En el primer caso, se aplicará el mismo método descrito anteriormente para determinar cuál es la siguiente variable que va a ser utilizada.

Así, las ramas se van dividiendo dando forma a un árbol, calculando en cada momento cuál es la mejor variable discriminatoria. Una rama habrá llegado a su última división cuando la ganancia ya no sea mayor que cero.

G.2.5. Técnica de «poda»

Un inconveniente que aparece frecuentemente cuando se entrena un árbol como el descrito anteriormente es que resulta ser muy específico para los datos de entrenamiento (está sobreajustado). Un árbol en estas condiciones puede dar como resultado una respuesta que es mucho más concreta de lo que en realidad debería ser, creando ramas que decremantan levemente la entropía para los datos de entrenamiento, pero cuyas condiciones son de hecho totalmente arbitrarias.

Dado que el algoritmo descrito anteriormente divide continuamente las ramas hasta que no se puede reducir más la entropía, una posibilidad es dejar de dividir cuando esta reducción es menor que una cantidad determinada. Esta estrategia se emplea con frecuencia, pero tiene un ligero inconveniente: es posible tener un conjunto donde la entropía no se reduce sustancialmente en una división pero lo hace de forma más significativa en las subdivisiones siguientes. Una estrategia alternativa es construir el árbol completo tal y como se ha descrito y después tratar de eliminar nodos superfluos. Este proceso es lo que frecuentemente se conoce como «poda» del árbol.

La poda implica comprobar pares de nodos que tengan un padre común para ver si mezclándolos se incrementaría la entropía en una cantidad menor que un valor umbral. Si es así, entonces las hojas se mezclan en un único nodo con todas las salidas posibles. Esta técnica ayuda a paliar el problema del sobreajuste y evita que el árbol realice predicciones más seguras de lo que se puede deducir a partir de los datos.

G.2.6. Trabajo con datos incompletos o perdidos

Otra ventaja en el empleo de los árboles de decisión es su capacidad para el trabajo con datos perdidos. El conjunto de datos puede carecer de un determinado fragmento de información. Para adaptar el árbol a estas condiciones, es necesario implementar una función de predicción diferente.

Si se carece de un determinado dato que se necesita para decidir cuál de las ramas del árbol deb seguirse, se pueden seguir al mismo tiempo ambas. No obstante, en lugar de contar los resultados de forma equitativa, deben ponderarse. En el árbol de decisión básico, todos los elementos tienen un peso implícito de 1, lo cual significa que las observaciones asumen la probabilidad de que un elemento encaje correctamente en una determinada categoría. Durante el seguimiento de múltiples

ramas del árbol, se les puede otorgar a cada una de ellas un peso igual a la fracción de todas las filas que están contenidas en esa parte.

G.2.7. Trabajo con salidas numéricas

Los clásicos ejemplos de clasificación se implementan para problemas en los que las salidas son categorías y no números como tales. Cuando es posible implementar un árbol sobre un conjunto de datos en los que las salidas son numéricas, probablemente el resultado no será del todo bueno. Si todos los números son tratados como categorías diferentes, el algoritmo no tendrá en cuenta el hecho de que algunos números son más cercanos y otros más lejanos; todos ellos serán tratados como números que están completamente separados. Para trabajar con esto, cuando se tiene un árbol con salidas numéricas puede utilizarse la varianza como métrica, en lugar de la entropía o la impureza de Gini.

Una varianza pequeña significa que los números son todos cercanos entre sí y una varianza elevada se traduce en que están ampliamente distribuidos. Cuando durante la construcción de un árbol se utiliza la varianza como métrica, los criterios en los nodos serán tales que los números quedarán divididos entre los que son mayores que uno en un lado y los inferiores en el otro. Dividir los datos de esta manera reduce la varianza global en las ramas.

G.2.8. Cuándo utilizar árboles de decisión

Probablemente la mayor ventaja de los árboles de decisión es lo sencillo que resulta interpretar un modelo entrenado. Después de ejecutar un algoritmo como el descrito anteriormente no sólo se finaliza con un árbol que puede hacer predicciones sobre nuevos datos, sino que también se dispone de una lista de preguntas que se utilizan para tomar decisiones. También se pueden conocer determinadas variables que en definitiva no son importantes a la hora de determinar la salida. Si la adquisición de los datos fuese difícil o costosa y se asume que no es importante, se sabe que se puede detener su recogida.

A diferencia de otros algoritmos de aprendizaje automático, los árboles de decisión pueden trabajar con datos de entrada en forma de categorías y de números. Además, mientras que otros algoritmos requieren un tratamiento previo de los datos para adaptarlos a un determinado formato, el algoritmo descrito anteriormente puede implementarse y ser ejecutado sin este requerimiento.

Los árboles de decisión también permiten la asignación probabilística de datos. En algunos problemas no existe suficiente información para hacer siempre una distinción correcta (un árbol de decisión puede tener un nodo que tiene varias posibilidades y no puede ser dividido más). El planteamiento aquí expuesto devuelve un diccionario de cuentas de diferentes salidas y esta información puede ser útil a la hora de decidir cómo de confidenciales son los resultados. No todos los algoritmos pueden estimar la probabilidad de un resultado incierto.

No obstante, hay algunos inconvenientes en los árboles de decisión planteados en este apartado. Mientras que pueden ser muy eficaces para problemas con solamente un pequeño conjunto de resultados, no pueden utilizarse de forma eficiente en conjuntos de datos con muchas posibilidades. En ese caso, el árbol de decisión crecería de forma muy compleja y probablemente tomaría decisiones de muy baja calidad.

La otra gran desventaja de estos árboles es que cuando pueden trabajar con datos numéricos simples, sólo se pueden crear decisiones del tipo mayor que o menor que. Esta cuestión hace difícil clasificar datos cuando la clase está definida por una combinación de variables más compleja. Por

ejemplo, si los resultados se determinan a partir de las diferencias de dos variables, el árbol crecería enormemente y se tornaría rápidamente impreciso.

En resumen, los árboles de decisión no son probablemente una buena elección para problemas con muchas entradas y salidas numéricas, o para aquellos que implican relaciones más complejas entre entradas numéricas. Los árboles son idóneos para conjuntos de datos de tipo categórico y datos numéricos que tengan puntos de ruptura. Estos árboles son la mejor elección si el entendimiento del proceso de decisión es importante: como se ha comentado anteriormente, el razonamiento puede llegar a ser tan importante como conocer la predicción final.

G.2.9. Detalles de implementación de árboles de decisión en la red UTRAN

Introducción

A lo largo de las páginas precedentes se han presentado los conceptos de inteligencia colectiva y aprendizaje automático. Particularizando dentro de esta última disciplina, se han expuesto los conocimientos generales sobre árboles de decisión, aspectos tales como qué son, cómo se construyen o cómo pueden implementarse.

El propósito de esta sección es aplicar la teoría vista anteriormente para el análisis de la accesibilidad en una red UTRAN. Así, se trata de construir un modelo de clasificación en el que a partir de la observación de los datos estadísticos reportados por los nodos B el sistema sea capaz de discriminar comportamientos anómalos singulares y tratar de predecir cuál puede ser la causa de los mismos, en base a cada uno de los factores limitantes que se han descrito a lo largo de la presente Tesis.

Definición de atributos

Como se ha visto anteriormente, una de las etapas críticas en el análisis de datos es la construcción del árbol de decisión. Para realizar su entrenamiento será necesario distinguir cuáles son las características de esos datos que serán objeto de análisis, cuáles las conclusiones a las que se deberá llegar y finalmente el diseño del árbol, que determinará el razonamiento a seguir para diagnosticar la posible causa de bloqueo.

Atributos de entrada y fuentes de datos

A partir de los datos de entrada, son las características que se estudiarán y sobre las que se interrogará para seguir el árbol. La fuente de los datos de entrada es el propio tráfico reportado por el nodo y la accesibilidad teórica reportada. A nivel temporal, la unidad básica de medida que se considera es un día completo.

Las características particulares de las redes móviles hacen en general inalcanzable un objetivo de accesibilidad del 100 %: factores físicos, sociales o económicos limitan las posibilidades de los operadores móviles a la hora de alcanzar este objetivo. Por ello, los operadores de red establecen un pequeño margen en el cual la accesibilidad puede fluctuar. Así, generalmente, se habla de un 99 % de accesibilidad como cota inferior para discriminar una anomalía. Consecuentemente, cada una de las estaciones tendrá un valor de accesibilidad media diaria que podría considerarse «habitual», entendiéndose que existe algún tipo de problema cuando los valores medidos se alejan de esta cantidad.



Figura G.2: Buffer o ventana de observación de accesibilidades reportadas

Considérese que los datos reportados de accesibilidad se almacenan en un buffer o ventana de observación de longitud «L», tal y como indica la figura G.2.

Cada uno de los valores almacenados corresponde a una observación diaria. Teniendo en cuenta que el nodo reporta datos agregados cada 15 minutos, se considera que el valor de accesibilidad diaria es el valor promedio de todas las muestras almacenadas, en este caso 96 elementos. La fórmula G.4 representa esta medida¹, que se almacena en cada una de las posiciones del buffer o ventana de observación.

$$\overline{Acc_{día}} = \frac{\sum_{i=1}^{96} CSSR(i)}{96} \quad (G.4)$$

Tal y como se ha comentado, es necesario determinar una referencia de cuál es la accesibilidad habitual, de modo que pueda ser utilizada como referencia. Para ello se empleará la media móvil de las accesibilidades, dejando a un lado las observaciones de los últimos 2 días para evitar el sesgo causado por los factores limitantes (ecuación G.5).

$$A_{m,n} = \frac{1}{m-n+1} \sum_{i=m}^n \overline{Acc_{día}} \quad (G.5)$$

A partir de dicha expresión, se definen dos métricas:

⇒ $A_{1,L-2}$ representa el valor histórico o habitual de la accesibilidad en la estación y es el dato de referencia.

⇒ $A_{L-1,L}$ son las últimas 2 muestras reportadas. La muestra «L-1» representa el «primer día anómalo» de manera que para descartar fallos puntuales se introduce en la métrica la última muestra.

Se define una constante de arbitraje «k» que representa la comparación de las 2 métricas anteriores, evaluando según la inequación G.6 si los dos últimos valores recogidos muestran un comportamiento anómalo.

$$\frac{A_{L-1,L}}{A_{1,L-2}} < k \quad (G.6)$$

Por tanto, si se satisface la inequación anterior se considerará que existe una incidencia y será entonces cuando será necesario evaluar cuál es el factor limitante en el sistema. Para ello, se utilizará

1

En las ecuaciones subsiguientes no se hará distinción entre accesibilidad de paquetes y de circuitos. Se entiende que estos cálculos se implementan de manera separada para ambos dominios, correspondiendo al lector la colocación de los subíndices correspondientes «CS» o «PS».

$CSSR$	$A_{L-1,L}$	$A_{1,L-2}$	$Acc_{teórica}$	Factor limitante
$CSSR_1$	Cálculo de parámetros			Iub
$CSSR_2$	Cálculo de parámetros			SF
$CSSR_3$	Cálculo de parámetros			CE
$CSSR_4$	Cálculo de parámetros			Iub
...
$CSSR_n$	Cálculo de parámetros			CE

Cuadro G.3: Formato y ejemplo de casos de entrenamiento para el árbol de decisión

	Conclusiones
# 1	Congestión en el enlace Iub.
# 2	Recursos de procesado insuficientes.
# 3	Códigos de canalización insuficientes.
# 4	Máxima potencia en el amplificador.
# 5	Elevado nivel de interferencia en el Uplink.
# 6	Fallo Hardware.

Cuadro G.4: Atributos de salida para el árbol de decisión

el análisis teórico presentado en capítulos anteriores, utilizando los datos de tráfico reportados por el nodo y modelando el sistema según el planteamiento descrito en capítulos previos.

Evaluando la expresión G.7 de forma positiva se concluye cuál es el factor limitante.

$$\frac{Acc_{teórica}}{Acc_x} > 1 \quad (G.7)$$

El entrenamiento del árbol de decisión se basa en la presentación y el contraste de datos de entrada y de salida, según los atributos definidos en los apartados precedentes. El cuadro G.3 representa, a través de una tabla de ejemplo, el formato de casos de entrenamiento que deben utilizarse.

Atributos de salida

Es el conjunto de conclusiones a las que deberá llegar el sistema de aprendizaje. En la tabla G.4 se recogen las distintas posibilidades².

G.3. Redes Neuronales

G.3.1. Concepto de Red Neuronal

Los desarrollos actuales de los científicos se dirigen al estudio de las capacidades humanas como una fuente de nuevas ideas para el diseño de las nuevas máquinas. Así, la inteligencia artificial es un intento por descubrir y describir aspectos de la inteligencia humana que pueden ser simulados mediante máquinas. Esta disciplina se ha desarrollado fuertemente en los últimos años teniendo aplicación en algunos campos como visión artificial, demostración de teoremas, procesamiento de información expresada mediante lenguajes humanos, etc.

Las redes neuronales son más que otra forma de emular ciertas características propias de los humanos, como la capacidad de memorizar y de asociar hechos. Si se examinan con atención

²Téngase en cuenta que las conclusiones #4 y #5 están planteadas a nivel teórico, puesto que no pueden implementarse con las estadísticas que recogen actualmente los Nodos B.

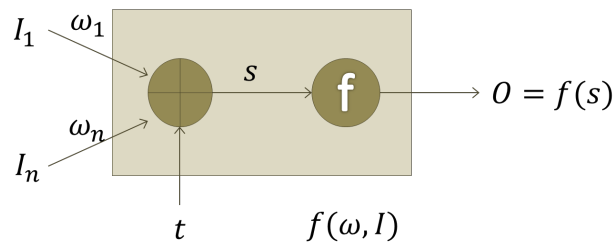


Figura G.3: Estructura básica de una neurona artificial

aquellos problemas que no pueden expresarse a través de un algoritmo, se observará que todos ellos tienen una característica en común: la experiencia. El hombre es capaz de resolver estas situaciones acudiendo a la experiencia acumulada. Así, parece claro que una forma de aproximarse al problema consista en la construcción de sistemas que sean capaces de reproducir esta característica humana.

Se entiende por red neuronal al sistema compuesto de un gran número de elementos básicos, denominados neuronas, agrupados en capas y altamente interconectados [96]. La red como tal estará formada por un conjunto de varias entradas y salidas, estas últimas entrenadas para reaccionar a los estímulos de entrada.

El desarrollo de una red neuronal implica dos actividades complementarias:

⇒ Aprender a comportarse (*Learning*).

⇒ Enseñarlas o entrenarlas (*Training*) en base a un conocimiento previo.

Los contextos en los que las redes neuronales resultan más útiles son aquellos en los que se dispone de una gran cantidad de datos y no se conoce con exactitud la estructura y los parámetros que modelan el problema.

Elemento básico: neurona artificial

Las neuronas son el elemento básico estructural de las redes neuronales. Son subsistemas formados por una o varias entradas y salidas, las cuales pueden ser binarias, analógicas o codificadas con pulsos, y constituyen la unidad básica de procesamiento. Se conectan con otras neuronas por medio de lo que se denominan conexiones sinápticas (ver figura G.6).

En la figura G.3 [96] puede verse un esquema general de la estructura de una neurona artificial. Está constituida por:

⇒ Entradas (I) ponderadas por unos pesos sinápticos ω_i .

⇒ Salida (O).

⇒ Función no lineal $f(s)$.

Los valores de las entradas ponderadas se suman y típicamente se comparan con un umbral t . El resultado de este sumatorio constituye la entrada de una función $f(s)$ no lineal, de manera que las expresiones G.8 y G.9 resumen el comportamiento de la neurona.

$$s = I_1\omega_1 + I_2\omega_2 + \dots + I_n\omega_n - t \quad (\text{G.8})$$

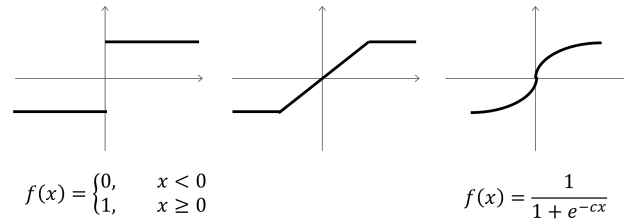


Figura G.4: Alternativas a la función $f(s)$ para una neurona artificial básica

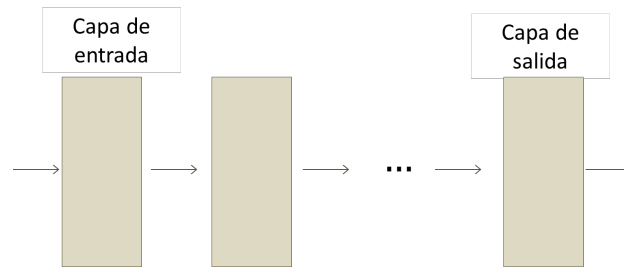


Figura G.5: Representación multicapa de una red neuronal

$$O = f(s) = f\left(\sum_{p=1}^n I_p \omega_p - t\right) \tag{G.9}$$

La elección de la función no lineal $f(s)$ determina el tipo de neurona y en parte el tipo de red neuronal. En la figura G.4 pueden verse los casos más típicos, respectivamente:

⇒ Función signo.

⇒ Función rampa.

⇒ Funciones sigmoidales. La expresión más típica es $f(x) = \frac{1}{1+e^{-cx}}$ y en ocasiones $f(x) = \tanh(cx)$.

Estructura de la red neuronal

La estructura de la red define cómo las neuronas comunican sus salidas a las entradas de otras neuronas. Típicamente, las neuronas se agrupan en capas, de modo que las salidas de una capa se conectan a las entradas de la siguiente, y así sucesivamente [96] (ver figuras G.5 y G.6). Para un resultado aceptable, es recomendable que el diseño de la red cuente al menos con tres capas.

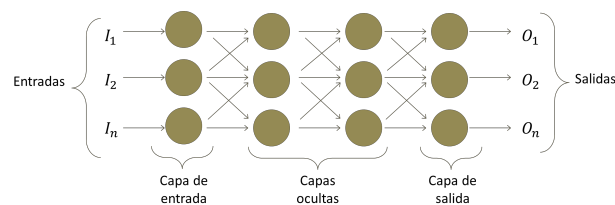


Figura G.6: Ejemplo de red neuronal multicapa con n entradas y n salidas

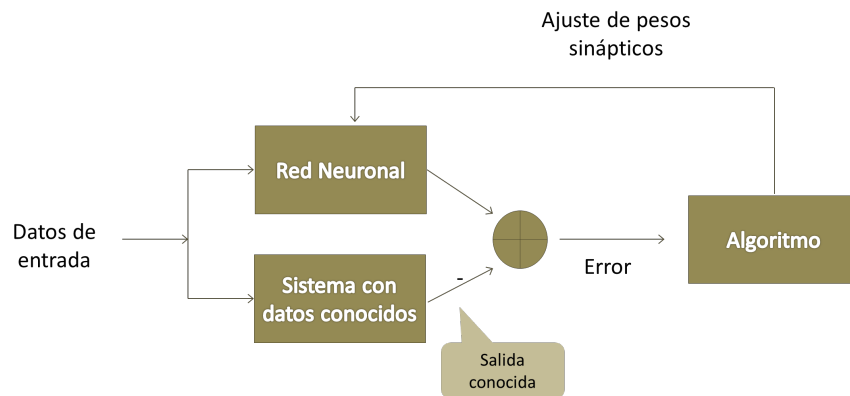


Figura G.7: Esquema básico de entrenamiento de una red neuronal

Aprendizaje. Proceso de entrenamiento

Este proceso comprende la adaptación progresiva de los valores de las conexiones sinápticas, para permitir a la red el aprendizaje de un comportamiento deseado (ver figura G.7) [96]. La metodología consiste en introducir los datos de entrada en la red neuronal, pero al mismo tiempo introducirlos también en un sistema con datos conocidos. En consecuencia, se dispondrá de dos datos de salida, uno el correspondiente a la red neuronal, y el otro ya conocido (es la salida del sistema conocido). A continuación se calcula una medida de calidad, y en función de este resultado se ejecuta un algoritmo que permite el ajuste de los pesos sinápticos.

Para la comprobación de la calidad del modelo pueden utilizarse diferentes métricas, típicamente una de las alternativas siguientes:

1. Suma de los cuadrados de los errores: $\sum_{p=1}^r \|b_p - \hat{b}_p\|^2$
2. Raíz cuadrada del error cuadrático medio: $\sqrt{\frac{1}{r} \sum_{p=1}^r \|b_p - \hat{b}_p\|^2}$
3. Error máximo: $\max \{ \|b_p - \hat{b}_p\|, p = 1 \dots r \}$

Ejecución

Una vez que se dispone de una red neuronal entrenada, puede alimentarse el sistema con una nueva entrada nunca antes vista, una situación nueva, y la red neuronal producirá entonces una respuesta razonable.

G.3.2. Aplicación de redes neuronales al modelado UTRAN

A lo largo de la presente Tesis se ha elaborado un modelo que emula en la medida de lo posible el comportamiento de un sistema real en la red UTRAN. El diseño ha partido inicialmente del encadenado de 3 subsistemas independientes con sus respectivos factores limitantes, aunque posteriormente se han ampliado a cinco a nivel descriptivo, ya que no resulta posible modelarlos al no tener estadísticos suficientes con la tecnología actual. En cualquier caso, en el capítulo previo se ha presentado un procedimiento de calibración, así como los resultados y la validez del modelo cuando éste se enfrenta a diferentes escenarios conocidos.

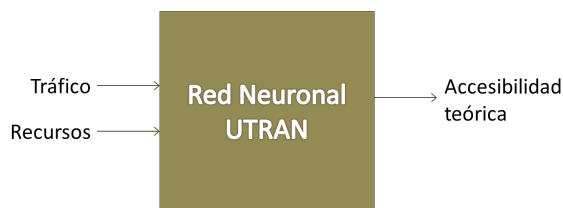


Figura G.8: Modelado de red UTRAN utilizando redes neuronales

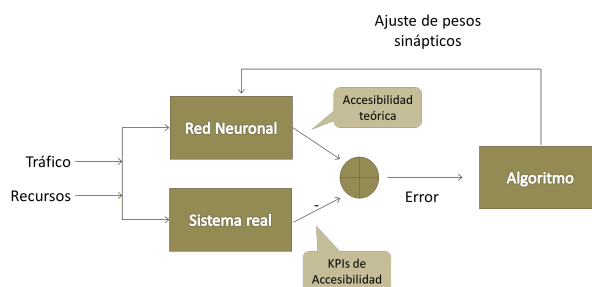


Figura G.9: Esquema de entrenamiento de red neuronal UTRAN

Queda pues patente que se ha conseguido realizar un modelado basándose en principios teóricos perfectamente conocidos (probabilidad de bloqueo en redes multiservicio con el algoritmo de Kaufman-Roberts), aunque algunos elementos del mismo se encuentren sujetos a diferentes grados de libertad (como la distribución del tráfico) y a pesar de que algunos de los subsistemas supuestos no han podido ser cuantificados. A esto hay que añadir que existen otros factores diversos que evidencian la complejidad que supone una reproducción fiel del modelo de red UTRAN.

En este contexto, y tal y como se ha comentado en el apartado anterior, las redes neuronales son una solución al modelado de sistemas en los que no se conoce con exactitud la estructura o la relación entre los diferentes parámetros del mismo. La propuesta es por tanto la sustitución del modelo anterior basado en la aplicación sucesiva del algoritmo de K-R por otro implementado mediante una red neuronal (ver figura G.8).

La estructura interna de la red neuronal como red multicapa excede el propósito de la presente Tesis, proponiéndose como futura línea de trabajo la implementación de la misma de acuerdo a las especificaciones aquí descritas.

Esquema de entrenamiento

De acuerdo con el esquema general presentado en la figura G.7, la red neuronal puede entrenarse ya que se parte de la ventaja de disponer de datos reales reportados directamente por las estaciones, como son el tráfico y la accesibilidad real. De este modo, considerando como parámetros de entrada el tráfico y los recursos de que se dispone, se introducen éstos en paralelo en la red neuronal y en el sistema real (el modelado del sistema real se dibuja sólo a título ilustrativo, pues las estadísticas ya están generadas). La red neuronal devolverá entonces la accesibilidad teórica, que será comparable al KPI de accesibilidad real. En función del error cometido, habrá de diseñarse un algoritmo que ajuste los pesos sinápticos de la red neuronal. En la figura G.9 se toma el procedimiento general de la figura G.7 y se particulariza para una red UTRAN.

G.4. Algoritmos genéticos

G.4.1. Concepto de algoritmo genético

Una vez que se ha planteado como alternativa el diseño de una red neuronal que modele con la mayor exactitud posible el sistema desarrollado en la Tesis, se tendrán al menos dos alternativas diferentes para un mismo modelo. Como cabe esperar, ambas soluciones darán resultados diferentes, cuyo grado de fidelidad habrá de ser posteriormente validado, de igual forma que en el capítulo anterior. Aquellos problemas así planteados en los que se tienen varias soluciones posibles, pueden estudiarse recurriendo a los llamados algoritmos genéticos.

Esta técnica se basa en los mecanismos de selección que utiliza la naturaleza, de acuerdo a los cuales los individuos más aptos de una población son los que sobreviven, al adaptarse más fácilmente a los cambios que se producen en su entorno. Hoy en día se sabe que estos cambios se efectúan en los genes de un individuo (unidad básica de codificación de cada uno de los atributos de un ser vivo), y que sus atributos más deseables (los que le permiten adaptarse mejor a su entorno) se transmiten a sus descendientes cuando éste se reproduce sexualmente. Los Algoritmos Genéticos (AGs) son métodos adaptativos que pueden usarse para resolver problemas de búsqueda y optimización [95]. Están basados en el proceso genético de los organismos vivos. A lo largo de las generaciones, las poblaciones evolucionan en la naturaleza de acorde con los principios de la selección natural y la supervivencia de los más fuertes, postulados por Darwin. Por imitación de este proceso, los Algoritmos Genéticos son capaces de ir creando soluciones para problemas del mundo real. La evolución de dichas soluciones hacia valores óptimos del problema depende en buena medida de una adecuada codificación de las mismas.

Un algoritmo genético consiste en una función matemática o una rutina de software que toma como entradas a los ejemplares y retorna como salidas cuales de ellos deben generar descendencia para la nueva generación [94]. Versiones más complejas de algoritmos genéticos generan un ciclo iterativo que directamente toma a la especie (el total de los ejemplares) y crea una nueva generación que reemplaza a la antigua una cantidad de veces determinada por su propio diseño. Una de sus características principales es la de ir perfeccionando su propia heurística en el proceso de ejecución, por lo que no requiere largos períodos de entrenamiento especializado por parte del ser humano, principal defecto de otros métodos para solucionar problemas, como los Sistemas Expertos.

G.4.2. Ventajas y desventajas. Limitaciones

- ⇒ No necesitan conocimientos específicos sobre el problema que intentan resolver.
- ⇒ Operan de forma simultánea con varias soluciones, en vez de trabajar de forma secuencial como las técnicas tradicionales.
- ⇒ Cuando se usan para problemas de optimización maximizar una función objetivo- resultan menos afectados por los máximos locales (falsas soluciones) que las técnicas tradicionales.
- ⇒ Resulta sumamente fácil ejecutarlos en las modernas arquitecturas masivamente paralelas.
- ⇒ Usan operadores probabilísticos, en vez de los típicos operadores determinísticos de las otras técnicas.
- ⇒ Pueden tardar mucho en converger, o no converger en absoluto, dependiendo en cierta medida de los parámetros que se utilicen tamaño de la población, número de generaciones, etc.

⇒ Pueden converger prematuramente debido a una serie de problemas de diversa índole.

Limitaciones

El poder de los Algoritmos Genéticos proviene del hecho de que se trata de una técnica robusta, y pueden tratar con éxito una gran variedad de problemas provenientes de diferentes áreas, incluyendo aquellos en los que otros métodos encuentran dificultades. Si bien no se garantiza que el Algoritmo Genético encuentre la solución óptima, del problema, existe evidencia empírica de que se encuentran soluciones de un nivel aceptable, en un tiempo competitivo con el resto de algoritmos de optimización combinatoria. En el caso de que existan técnicas especializadas para resolver un determinado problema, lo más probable es que superen al Algoritmo Genético, tanto en rapidez como en eficacia. El gran campo de aplicación de los Algoritmos Genéticos se relaciona con aquellos problemas para los cuales no existen técnicas especializadas. Incluso en el caso en que dichas técnicas existan, y funcionen bien, pueden efectuarse mejoras de las mismas hibridándolas con los Algoritmos Genéticos.

G.4.3. Funcionamiento de los algoritmos genéticos

Por analogía con la disciplina de la biología, se parte de un conjunto de soluciones a un problema, que de ahora en adelante será llamado población de individuos. A cada uno de estos individuos se le otorga una puntuación, de acuerdo a la bondad de la solución (que puede ser medida con distintas métricas). A mayor puntuación, significa que la probabilidad de seleccionarse para un cruce será mayor [94].

Se denomina cruce al proceso por el cual se generan nuevos individuos llamados descendientes, que comparten características de los padres. Es decir, soluciones al problema que contienen una mayor proporción de buenas características que la generación anterior. De esta manera, la población converge hacia una solución óptima.

Algoritmo genético simple

A continuación se resume la metodología más sencilla a la hora de implementar un algoritmo genético [94, 95]. En primer lugar, se enumeran las diferentes etapas y posteriormente en párrafos inferiores se profundiza en algunas de ellas:

1. Selección de dos individuos del conjunto de la población (selección de dos soluciones al problema).
2. Cruzar los dos individuos generando dos descendientes (generación de dos nuevas soluciones al problema que reemplazan a las anteriores).
3. Mutar los descendientes. El proceso de mutación consiste en efectuar pequeñas variaciones con una probabilidad muy baja.
4. Evaluar los descendientes mutados (evaluar la bondad de las nuevas soluciones del problema).
5. Insertarlos en la nueva generación y repetir el proceso cuantas veces sea necesario.

Puede considerarse que los individuos están formados por un conjunto de parámetros (en adelante genes). Cada uno de los genes está formado por cromosomas, y la definición de un gen como

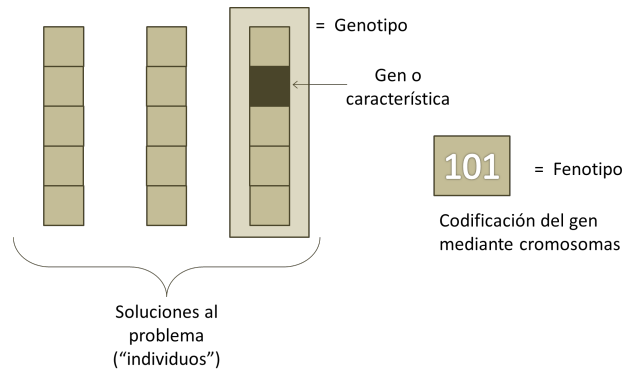


Figura G.10: Definición general de los individuos de un algoritmo genético

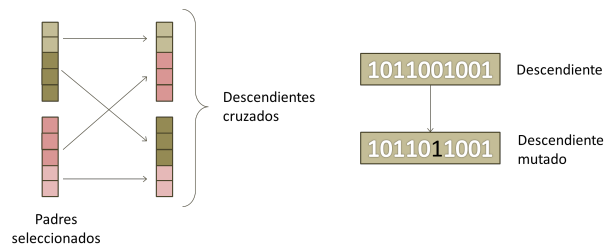


Figura G.11: Operaciones de cruce de poblaciones y mutación de descendientes

vector de cromosomas se denomina fenotipo (figura G.10). La función que evaluará la bondad de la solución será una función de evaluación del fenotipo y representa la adaptación del individuo al problema.

La selección de los padres se efectúa en principio al azar, con un procedimiento que favorezca a los individuos mejor adaptados, un proceso que se conoce como *ruleta degradada* [94]. Cada individuo (o cada solución al problema) tiene una probabilidad de ser seleccionado.

Una vez que se ha seleccionado a los padres, se efectúa sobre ellos dos operaciones: cruce y mutación (figura G.11).

⇒ La operación de cruce consiste en fragmentar la solución en subconjuntos de varios genes e intercambiarlos. De esta forma, los descendientes toman genes de ambos padres, o lo que es lo mismo, la solución al problema tendrá características de los dos problemas previos. Este proceso converge a una solución mejor, puesto que si la selección está bien perfeccionada los padres son los mejor adaptados de su generación. Esta operación de cruce sólo se hace con una probabilidad acotada entre 0,5 y 1, puesto que en caso contrario se podría incurrir en una simple duplicidad de los padres y el algoritmo no evolucionaría.

⇒ A continuación se efectúa la operación de mutación, que se aplica de manera individual a cada individuo. Consiste en alterar, con una probabilidad muy pequeña, cada gen o característica de la solución.

Se dice que una población converge cuando convergen todos los genes que la forman, y se dice que un gen converge cuando más del 95% de los individuos comparten ese valor del gen. En tal caso se dice que el algoritmo genético ha evolucionado hacia una generación formada por soluciones al problema con mayor bondad.

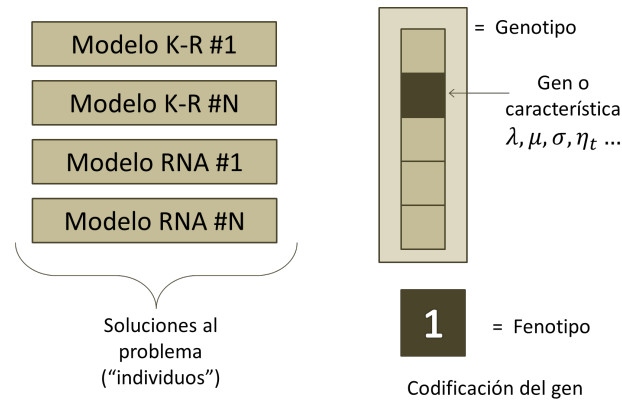


Figura G.12: Aplicación de algoritmos genéticos para modelado UTRAN

G.4.4. Aplicación de los algoritmos genéticos al modelado UTRAN

Partiendo de la base de que el modelado de la red UTRAN se ha realizado desde varios puntos de vista, a saber:

- ⇒ Modelado basado en desarrollos teóricos según algoritmo multiservicio de Kaufman-Roberts. Es un modelo en el que se han tenido en cuenta 3 factores limitantes, sabida la existencia de más factores que pueden modificar el resultado. Además, es un modelo que admite diferentes soluciones de acuerdo a variables intrínsecas al modelo: interpolación y calibración del tráfico, ponderación de accesibilidades parciales, medida de calidad, umbral de digitalización, etc. La combinación de varias de estas alternativas da como resultado diferentes soluciones al problema. En algunos casos se producen mejoras y en otras se alcanza un máximo de bondad, que la calibración del modelo no es capaz de mejorar.
- ⇒ Modelado basado en redes neuronales. En función de la implementación de la red se obtendrán diferentes soluciones, cuya fiabilidad habrá de ser evaluada.

Resta decir que, al intervenir diferentes factores físicos en el problema (en términos de propagación y de las condiciones específicas de cada escenario) no resulta sencillo extrapolar un único modelo a diferentes escenarios, ya que habría que considerar un número incontable de factores difícilmente cuantificables.

En resumen, se tienen diferentes soluciones al problema, provenientes de modelos de naturaleza muy diferente. La evaluación de todos ellos da como resultado diferentes soluciones en diferentes escenarios, habida cuenta de que todas ellas presentan un grado de similitud aceptable. Por ello, una vez que se tienen diferentes soluciones al problema, es posible la aplicación de los algoritmos genéticos que proporcionen nuevas soluciones más adaptadas a este contexto. En la figura G.12 se plantea lo que puede ser una abstracción del modelo de red UTRAN propuesto en anteriores apartados y la posible aplicación posterior de los algoritmos genéticos.

Apéndice H

Material empleado

H.1. Material empleado

Para la elaboración de esta Tesis, además de las herramientas comentadas en el punto 7.11: Opnet y MATLAB para el modelado del Iub, se han utilizado otras herramientas software que se destacan a continuación:

- ⇒ El proyecto ha sido desarrollado indistintamente sobre plataformas MacOS X Snow Leopard 10.6, GNU/Linux (distribución Ubuntu 10.10 Maverick Meerkat) y Windows 7 (por orden de tiempo empleado).
- ⇒ Para la edición de imágenes vectoriales se han utilizado las herramientas Adobe Illustrator CS5 y Corel Draw X4.
- ⇒ Para la edición de imágenes rasterizadas se ha empleado Adobe Photoshop CS5.
- ⇒ En la redacción del documento se ha empleado el entorno Latex mediante las distribuciones MacTeX, T_EX Live y MikTeX y mediante compilaciones PDFL_AT_EX. Como entorno de desarrollo Latex se ha decidido emplear el programa Lyx, dado su carácter multiplataforma.
- ⇒ Opnet 14.5.2 con licencia de la Universidad de Oviedo.
- ⇒ Matlab R2010b.

Glosario de acrónimos

3GPP: *3rd Generation Partnership Project.*

A

AAAL: *ATM Adaption Layer.*

ADM: *Adaptative Delta Modulation.*

ADSL: *Asymmetric Digital Subscriber Line.*

AMR: *Adaptative Multi-Rate.*

AMS: *American Phone System.*

ANSI: *American National Standards Institute.*

ATM: *Asynchronous Transfer Mode.*

AuC: *Autentication Center.*

B

BSC: *Base Station Controller.*

BSS: *Base Station Subsystem.*

BTS: *Base Transceiver Station.*

BW: *BandWidth.*

C

CAC: *Call Admission Control.*

CBR: *Constant Bit Rate.*

CBU: *Control Base Unit.*

CDMA: *Code Division Multiple Access.*

CPS: *Common Part Sub-layer.*

CRC: *Cyclic Redundancy Check.*

CE: *Channel Element.*

CS: *Circuit Switched.*

CS: *Convergence SubLayer.*

D

DCH: *Dedicated transport CHannel.*

DCS: *Digital Cellular System.*

DL: *DownLink.*

DSCH: *Downlink Shared CHannel.*

DSL: *Digital Subscriber Line.*

DTX: *Discontinuous Transmission.*

E

EcNo: *Relación señal a interferencia a nivel de bit.*

EIR: *Equipment Identity Register.*

ETB: *Exchange Terminal Board.*

F

FACH: *Forward Access CHannel.*

FP: *Framing Protocol.*

G

GERAN: *GSM/Edge Radio Access Network.*

GGSN: *Gateway GPRS Support Node.*

GMSC: *Gateway MSC.*

GoS: *Grade of Service.*

GPRS: *General Packet Radio Service.*

GPRS-CN: *General Packet Radio Service, Core Network.*

GSM: *Global System for Mobile Communications.*

GTP: *GPRS Tunelling Protocol.*

GUI: *Graphical User Interface.*

H

HLR: *Home Location Register.*

HSPA: *High-Speed Packet Access.*

HSDPA: *High Speed Downlink Packet Access.*

HSUPA: *High-Speed Uplink Packet Access.*

I

IDE: *Integrated Development Environment.*

IMA: *Inverse Multiplexing over ATM.*

IMT-2000: *International Mobile Telecommunications-2000.*

IP: *Internet Protocol.*

ITU: *International Telecommunication Union.*

K

KPI: *Key Performance Indicator.*

M

MAC: *Media Access Control.*

MCR: *Minimum Cell Rate.*

MIMO: *Multiple-Input Multiple-Output.*

MPLS: *Multiprotocol Label Switching.*

MS: *Mobile Station.*

MSC: *Mobile Switching Center.*

N

NBAP: *Node B Application Part.*

NMT: *Nordic Mobile Telephone.*

NRT: *Near Real Time.*

NSS: *Network and Switching Subsystem.*

O

OBIF: *Optical Radio Unit Interface Board.*

OSI: *Open System Interconnection.*

P

PCM: *Pulse Code Modulation.*

PCR: *Peak Cell Rate.*

PCU: *Packet Control Unit.*

PDH: *Plesiochronous Digital Hierarchy.*

PDU: *Protocol Data Unit.*

PS: *Packet Switched.*

PSTN: *Public Switched Telephone Network.*

PTI: *Payload Type Indicator.*

PTN: *Packet Transport Network.*

Q

QoS: *Quality Of Service.*

R

R99: *Release 99.*

RAB: *Radio Access Bearer.*

RACH: *Random Access Channel.*

RANAP: *Radio Access Network Application Part.*

RAXB: *Base band Random Access Receiver Board.*

RB: *Radio Bearer.*

RBS: *Radio Bearer System.*

RLC: *Radio Link Control.*

RNC: *Radio Network Controller.*

RNS: *Radio Network Subsystem.*

RRC: *Radio Resource Control.*

RRU: *Remote Radio Unit.*

S

SAR: *Re-Assembly Sublayer.*

SC: *Spreading Code.*

SDH: *Synchronous Digital Hierarchy.*

SDU: *Service Data Unit.*

SF: *Spreading Factor.*

SDI: *Silence Insertion Descriptor.*

SGSN: *Serving GPRS Support Node.*

SHO: *Soft HandOver.*

SONET: *Standard for Optical NETWORKS.*

SRB: *Signalling Radio Bearer.*

SSCS: *Service Specific Convergence Sub-layer.*

STM: *Synchronous Transport Module.*

T

TB: *Transport Block.*

TDM: *Time Division Multiplexing.*

TFI: *Transport Format Indicator.*

TTI: *Transmission Time Interval.*

TU: *Tributary Unit.*

TUG: *Tributary Unit Group.*

TXB: *Base band Transmitter Board.*

U

UBR: *Unspecified Bit Rate.*

UL: *UpLink.*

UMTS: *Universal Mobile Telecommunications System.*

UTRAN: *UMTS Terrestrial Radio Access Network.*

V

VBA: Visual Basic for Applications.

VCC: *Virtual Channel Connection.*

VLR: *Visitor Location Register.*

VCI: *Virtual Channel Identifier (ATM).*

VPI: *Virtual Path Identifier (ATM).*

W

WCDMA: *Wideband Code Division Multiple Access.*

WFQ: *Weighted Fair Queuing.*

WiMAX: *Worldwide Interoperability for Microwave Access.*

X

xDSL: *DSL Technologies.*

Índice alfabético

- árbol, 294
- AAL1, 45, 237
- AAL2, 45, 238, 289
- AAL3-4, 45, 238
- AAL5, 45, 238
- ABR, 44, 237
- algoritmo genético, 307
- Aprendizaje automático, 294
- ATM, 42, 235, 289
- ATM, adaptación, 45, 237
- ATM, celdas, 43, 235
- ATM, conexiones virtuales, 43, 236
- ATM, conmutación, 43, 236
- ATM, QoS, 44, 236
- ATM, señalización, 243
- AuC, 22
- Background, 72
- Backhaul, 42
- BSC, 22
- BSS, 22
- BTS, 22
- códigos, árbol, 88
- códigos, ensanchamiento, 88
- CBR, 44, 237
- cruce, 309
- DCS, 19
- DualStack, 289
- EIR, 22
- Entorno, 139
- entrenamiento, 305
- Erlang B, 60
- estadísticas, procesado, 137
- Ethernet, radioenlace, 41
- Fibra Óptica, 42
- flujos garantizados, 76
- flujos no garantizados, 76
- Full-IP, 290
- GERAN, 22
- GGSN, 22
- GPRS, 22
- GPRS-CN, 22
- GSM, 19
- GSM, arquitectura, 21
- HLR, 22
- HSPA, algoritmo, 249
- HSPA, modelado, 87
- IMT-2000, 20
- Inteligencia Colectiva, 293
- Interactive, 72
- Iub, dimensionamiento IP, 96
- Java, 142
- KPIs, 283
- Material, 311
- MATLAB, 143
- modulación adaptativa, 41
- MSC, 22
- mutación, 309
- neurona, 303
- Nodo B, 30
- NSS, 22
- Opnet, 139
- PCU, 22
- PDH, 239
- PDH, radioenlace, 41
- Python, 142

RAB, 70
Radio Bearer, 70
Radioenlace, 40
RB, 70
RBS, 30
Recursos, 83
Red Neuronal, 302
RNC, 30

SC, 88
SDH, 240
SDH, radioenlace, 41
Servicios Portadores, 69
SONET, 240
STM-1, 240
Streaming, 72

Tráfico, clases, 71
tráficos elásticos, gestión, 84
tráficos rígidos, gestión, 84

UBR, 44, 237
UMTS, 20
UTRAN, arquitectura, 30

VBR-NRT, 44, 237
VBR-RT, 44, 237
Visual Basic, 141
VLR, 22
Vocoder, 245

WCDMA, 87
WFQ, 87

xDSL, 41

Bibliografía

Trabajos previos y estado del arte

- [1] Xavier Pérez-Costa, Albert Branchs, Kjell Heinze. Analysis of performance issues in an ip-based umts radio access network. *International Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems - MSWiM*, 2005.
- [2] R. Nunez Queija, J.L. van den Berg, M.R.H. Mandjes. Performance evaluation of strategies for integration of elastic and stream traffic. *Proceedings of International Teletraffic Congress*, 1999.
- [3] Goran Eneroth, Martin Jonsson. Atm transport in cellular networks. *Proc. Of the International Switching Symposium, ISS 97, Toronto*, 1997.
- [4] IP Transport in UTRAN Technical Specification Group Radio Access Network. *3GPP TR 25.993*, 2003.
- [5] Ip in the ran as a transport option in 3rd generation mobile systems. *Mobile, Wireless Internet Forum, MTR-006, v2.0.0*, 2001.
- [6] Lieve Bos, Suresh Leroy. Toward an all-ip-based umts system architecture. *IEEE Network*, 2001.
- [7] Jin Yang, Ioannis Kriaras. Migration to all-ip based umts networks. *3G Mobile Communication Technologies, 2000. First International Conference on (Conf. Publ. No. 471)*, 2000.
- [8] D. Hervé A. Zwi K. Rana, M. Lars, B. Raquel. A bayesian approach for automated troubleshooting for umts networks. *Personal, Indoor and Mobile Radio Communications, 2006 IEEE 17th International Symposium on*, strony 1–5, 2006.
- [9] M. Stasiak, J. Wiewióra. Analytical model of the radio interface in the umts with packet scheduling mechanism. *AFRICON, 2009*, strony 1–6, 2009.
- [10] Andreas Mader Dirk Staehle. An analytic approximation of the uplink capacity in a umts network with heterogeneous traffic. *Proc. 18th International Teletraffic Congress (ICT), Berlin*, strony 81–91, 2003.
- [11] A. Mäder, D. Staehle. Analytic modeling of the wcdma downlink capacity in multi-service environments. *16th ITC Specialist Seminar on Performance Evaluation of Wireless and Mobile Systems*, strony 217–226, 2004.
- [12] G. Fodor, M. Telek. A recursive formula to calculate the steady state of cdma networks. *Proc. ITC 19, Beijing, China*, strony 1285–1294, 2005.

- [13] G. Fodor, M. Telek. Bounding the blocking probabilities in multirate cdma networks supporting elastic services. *IEEE/ACM Transactions on Networking (TON)*, 15:944–956, 2007.
- [14] G. Fodor, M. Telek. Performance analysis of the uplink of a cdma cell supporting elastic services. *NETWORKING 2005. Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems*, 3462:123–149, 2005.
- [15] Jerzy Martyna. A new method for computing blocking probabilities in cellular mobile communication networks. *Theoretical and Applied Informatics*, 22:165–175, 2010.
- [16] M. Hosein Fallah Y. Ouyang. A study of throughput for iu-cs and iu-ps interface in umts core network. *Performance Computing and Communications Conference (IPCCC)*, strony 349–353, 2009.
- [17] B. Jones X. Yan, J. Y Khan. Impact of a radio access network capacity on the hsdpa link performance. *Vehicular Technology Conference IEEE 66th*, strony 220–224, 2007.
- [18] P. Zwierzykowski M. Stasiak, J. Wiewiora. Model of iub interface in the umts network. *Ubiquitous Computing and Communication Journal*, strony 1–7, 2010.
- [19] J. Triola R. Barco L. Moltsen Z. Altman P. Lazaro R. M Khanafer, B. Solana. Automated diagnosis for umts networks using bayesian network approach. *Vehicular Technology, IEEE Transactions on*, strony 2451–2461, 2008.
- [20] Y.J. Chung H.G. Kim, C.H. Paik. Heuristics for the access network design problem in 3g mobile communication networks. *Innovative Computing Information and Control, 2008. ICICIC 08. 3rd International Conference on*, strony 254–254, 2008.
- [21] J. Sanchez Gonzalez R. Agusti Comes M.A. Diaz Guerra D. Henche D. Paul O. Sallent, J. Perez Romero. A roadmap from umts optimization to lte self-optimization. *IEEE communications magazine*, 6:172–182, 2011.
- [22] V Pla Martinez Bauset D. Garcia Roger, M Domenech Benlloch. Maximizing the capacity of mobile cellular networks with heterogeneous traffic. *Computer Networks*, 2008.
- [23] S. Tabbane A. Masmoudi. Optimized dimensioning methods for hspa based beyond 3g mobile networks. *Information Infrastructure Symposium, 2009. GIIS 09. Global*, strony 1–8, 2009.
- [24] M. Lohmann C. Lindemann. Traffic modeling and characterization for umts networks. *GLOBECOM*, strony 1741–1746, 2001.
- [25] P. Zwierzykowski M. Stasiak, J. Wiewióra. Analytical model of the wcdma radio interface in umts network with user differentiation. *2009 Fifth Advanced International Conference on Telecommunications*, strony 83–88, 2009.
- [26] J. Kaufman. Blocking in a shared resource environment. *Communications, IEEE Transactions on*, 29:1474–1481, 1981.
- [27] R. Kreher. *UMTS performance measurement: a practical guide to KPIs for the UTRAN environment*. Wiley, 2006.

- [28] J. Wiewióra M. Stasiak, P. Zwierzykowski. Analytical modelling of the iub interface in the umts network. *Communication Systems, Networks and Digital Signal Processing, 2008. CNSDSP 2008. 6th International Symposium on*, strony 204–207, 2008.
- [29] P. Zwierzykowski M. Stasiak, D. Parniewicz. Modelling of the wcdma interface in the umts network with soft handoff mechanism. *Global Telecommunications Conference, GLOBECOM. IEEE*, strony 1–6, 2009.
- [30] P. Zwierzykowski M. Stasiak, J. Wiewiora. Wcdma interface in umts network carrying a mixture of multi-rate traffic with different priorities. *Telecommunications, 2009. ICT 09. International Conference on*, strony 254–259, 2009.
- [31] P. Zwierzykowski M. Stasiak, J. Wiewióra. Modeling and dimensioning of the iub interface in the umts network. *Ubiquitous Computing and Communication Journal. Special Issue CSNDSP*, 2009.
- [32] J.-M. Picard B. Fourestié S. B. Jamaa, Z. Altman, J. Mourlon. Manual and automatic design for umts networks. *Mobile Networks and Applications*, 9:619–626, 2004.
- [33] C. Goerg X. Li, R. Schelb, A. Timm-Giel. Dimensioning of utran iub links for elastic internet traffic with multiple radio bearers. *Measuring, Modelling and Evaluation of Computer and Communication Systems (MMB)*, strony 1–16, 2006.
- [34] C. Gorg X. Li, R. Schelb, A. Timm-Giel. Dimensioning of utran iub links for elastic internet traffic. 2005.
- [35] Salman AlQahtani, Ashraf S. Mhmoud. A qos-aware call admission control algorithm for 3g cellular wireless networks. *King Fahd University of Petroleum and Minerals, Dhahran , Saudi Arabia. GCC Conference (GCC)*, 2006.
- [36] M. Andreas, S. Dirk. Spatial and temporal fairness in heterogeneous hsdpa-enabled umts networks. *EURASIP Journal on Wireless Communications and Networking*, 2009, 2009.
- [37] Antonio Capone, Simone Redana. Performance of call admission control techniques for td-cdma systems. *DEI, Politecnico di Milano, Vehicular Technology Conference IEEE 58th*, 2003.
- [38] Ramón Agustí Comes, Oriol Sallent roig, Manuel Álvarez Campana Fernández-Corredor. *Elementos de arquitectura y gestión de recursos radio en UMTS*. Fundación Vodafone España, 2005.
- [39] Álvaro Pachón de la Cruz. Evolución de los sistemas móviles celulares gsm. *Departamento Redes y Comunicaciones Universidad Icesi-I2T*, 2004.
- [40] A. Wacker D. Soldani, K. Sipila. Provisioning radio access networks for effective qos management: capacity gains of service differentiation in utran. *World of Wireless Mobile and Multimedia Networks, 2005. WoWMoM 2005. Sixth IEEE International Symposium on a*, strony 484–486, 2005.
- [41] Khedher F. Valois, S. Tabbane. Traffic characterization for mobile networks. *Proceedings IEEE 56th Vehicular Technology Conference*, 00-C:1485–1489, 2002.

- [42] Janos Farkas, Csaba Antal, Sandor Racz. Dimensioning link capacity in a packet switched telecommunications network. *World Intellectual Property Organization nº WO 2008/120077 A2*, 2008.
- [43] Gabor Fodor, Gosta Leijonhufvud, Szabolcs Malomsoky. Comparison of call admission control algorithms in atm/aal2 based 3rd generation mobile access networks. *Traffic Analysis and Network Performance Laboratory, Ericsson Research*.
- [44] A. B. García, M. Álvarez Campana, E. Vázquez. Diseño de redes de acceso en sistemas móviles umts con soporte de calidad de servicio. 2002.
- [45] Vijay K. Garg, Oliver T.W. Yu. Integrated qos support in 3g umts networks. *Wireless Communications and Networking Conference*, 2000.
- [46] V. B. Iversen. Evaluation of multi-service cdma networks with soft blocking. *Smart Spaces and Next Generation Wired/Wireless Networking*, 6294:160–171, 2010.
- [47] V. B. Iversen. Modelling restricted accessibility for wireless multi-service systems. *Wireless Systems and Network Architectures in Next Generation Internet*, 3883:93–102, 2006.
- [48] T.E. Kolding K.I. Pedersen, P.E. Mogensen. Qos considerations for hsdpa and performance results for different services. *Vehicular Technology Conference, VTC IEEE 64th*, strony 1–5, 2006.
- [49] László Korössi, Csaba Vulkán. Qos aware hsdpa congestion control algorithm. *Networking and Communications, 2008. WIMOB 08. IEEE International Conference on Wireless and Mobile Computing*, 2008.
- [50] Julio César Escobar de León, Herbert Alonso Solórzano. Servicios multimedia de tercera generación de telefonía móvil con el estándar umts. *Escuela de Ingeniería en Ciencias y Sistemas, Universidad San Carlos de Guatemala*, 2005.
- [51] I. D Moscholios M. Logothetis, V. G Vassilakis. Call-level performance modeling and qos assessment of w-cdma networks. *Wireless Networks: Research, Technology and Applications*, strony 57–90, 2009.
- [52] E. Vivier M. Terré, B. Fino. Optimisation of downlink resource allocation algorithms for umts networks. *EURASIP J. Wirel. Commun. Netw*, 2005:573–578, 2005.
- [53] Matthias Malkowski, Michael Schnick, Marc Schinnenburg. Connection admission control in umts with respect to network capacity and quality of service. *Wireless Conference 2005 - Next Generation Wireless and Mobile Communications and Services (European Wireless), 11th European*, 2005.
- [54] Szabolcs Malomsoky, Szilvester Nádas, Balázs Sonkoly. Performance evaluation of umts terrestrial radio access networks. *Traffic Analysis and Network Performance Laboratory, Ericsson Research*.
- [55] Szabolcs Malomsoky, Sandor Racz, Szilveszter Nadas. Connection admission control in umts radio access networks. *Traffic Analysis and Network Performance Laboratory, Ericsson Research*, 2003.

- [56] New ITU-T Rec. Q2630. Aal type 2 signaling protocol. 2000.
- [57] Hervé Dubreil Zwi Altman Rana Khanafer, Lars Moltsen, Raquel Barco. A bayesian approach for automated troubleshooting for umts networks. *The 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 06)*, stromy 1–6, 2005.
- [58] J. W. Roberts. A service system with heterogeneous user requirements application to multi-service telecommunications systems. g. pujolle, redaktor. *Proceedings of Performance of Data Communications Systems and their Applications*, stromy 423–431, 1981.
- [59] D. Rui y otros S. Nuno, S. Armando. Multiresolution with hierarchical modulations for long term evolution of umts. *EURASIP Journal on Wireless Communications and Networking*, 2009, 2009.
- [60] D. Zhang S.J. Oh. Optimal resource allocation in multiservice cdma networks. *Wireless Communications, IEEE Transactions on*, 2:811–821, 2003.
- [61] E. M. Saad, O. M. EL-Ghandou, M. K Jehan. Evaluation of qos in umts backbone network using differentiated services. *Radio Science Conference NRSC*, 2008.
- [62] Jagdeep Singh. Performance aspects of enhanced radio resource management with qos differentiation in umts network. *3G and Beyond, 2005 6th IEE International Conference On*, 2005.
- [63] Carlos R. Storck, Anna Izabel J. T. Ribeiro, Fátima L. P. Duarte-Figueiredo. Cac-rd: A call admission control for umts networks. *Computer Communications and Networks, 2008. ICCCN 08. Proceedings of 17th International Conference on*, 2008.
- [64] Cynthia Castellano Talavera, Jose María Hernando Rábanos. Marco de qos para aplicaciones de datos en umts. *Departamento de Planificación y Optimización Radio Telefónica Móviles España, Departamento de Señales Sistemas y Radiocomunicaciones Univesidad Politécnica de Madrid*, 2005.
- [65] L. Popova V. Benetis, V. B. Iversen. Joint connection and packet level analysis in w-cdma radio interface. *Wireless Systems and Mobility in Next Generation Internet*, 4396:186–199, 2007.
- [66] I. D. Moscholios V. G. Vassilakis, G. A. Kallos, D. M. Logothetis. The wireless engset multi-rate loss model for the call-level analysis of w-cdma networks. *Personal, Indoor and Mobile Radio Communications. IEEE 18th International Symposium on*, stromy 1–5, 2007.
- [67] I. D. Moscholios V. G. Vassilakis, G. A. Kallos, M. D. Logothetis. Call-level analysis of w-cdma networks supporting elastic services of finite population. *Communications, 2008. ICC 08. IEEE International Conference on*, stromy 285–290, 2008.
- [68] I. D. Moscholios V. G. Vassilakis, G. A. Kallos, M. D. Logothetis. On the handoff-call blocking probability calculation in w-cdma cellular networks. *Telecommunications, 2008. AICT 08. Fourth Advanced International Conference on*, stromy 173–179, 2008.

- [69] G. Hadjipollas A. Pitsillides V. Vassiliou, J. Antoniou. A simulation tool to evaluate radio resource management algorithms for enhanced umts. *2005. WIOPT 2005. Third International Symposium on*, stromy 396–403, 2005.
- [70] K. Venken, D. De Vleeschauwer, J. De Vriendt. Dimensioning the backbone network for an ip-based utran. *Proceedings of the 2001 International Conference on Third Generation Wireless and Beyond*, 2001.
- [71] T. Weerawardane. *Optimization and Performance Analysis of High Speed Mobile Access Network*. Vieweg and Teubner, 2011.
- [72] C. Gorg A. Timm-Giel X. Li, S. Li. Traffic modeling and characterization for utran. *Wireless/Wireless Internet Communications*, stromy 190–201, 2006.
- [73] Y. Ouyang, M. H. Fallah. A performance analysis for umts packet switched network based on multivariate kpis. *International Journal of Next-Generation Networks 2.1*, stromy 79–92, 2010.
- [74] Physical layer aspects of ultra high speed downlink packet access. *3GPP TR 25.848 v4.0.0*, 2000.

Métodos de optimización y dimensionamiento de recursos

- [75] J.M. Picard B. Fourestié y J. Mourlon S.B. Jamaa, Z. Altman. Manual and automatic design for umts networks. *Mobile Networks and Applications*, 9:619–626, 2004.
- [76] J. Wiewiora M. Stasiak. Analytical model of the radio interface in the umts with packet scheduling mechanism. *AFRICON*, pages 1–6, 2009.
- [77] y J. Wiewiora M. Stasiak, P. Zwierzykowski. Analytical model of the wcdma radio interface in umts network with user differentiation. *2009 Fifth Advanced International Conference on Telecommunications*, pages 83–88, 2009.
- [78] y J. Wiewiora M. Stasiak, P. Zwierzykowski. Wcdma interface in umts network carrying a mixture of multi-rate traffic with different priorities. *Telecommunications, 2009. ICT 09. International Conference on*, pages 254–259, 2009.
- [79] C. Goerg y A. Timm Giel X. Li, R. Schelb. Dimensioning of utran iub links for elastic internet traffic. *International Teletraffic Congress, Beijing*, 2005.
- [80] C. Goerg y A. Timm Giel X. Li, R. Schelb. Dimensioning of utran iub links for elastic internet traffic with multiple radio bearers. *Measuring, Modelling and Evaluation of Computer and Communication Systems (MMB), 2006 13th GI/ITG Conference*, pages 1–16, 2006.
- [81] P. Zwierzykowski M. Stasiak and D. Parniewicz. Modelling of the wcdma interface in the umts network with soft handoff mechanism. *Global Telecommunications Conference, 2009. GLOBECOM. IEEE*, pages 1–6, 2009.
- [82] y P. Zwierzykowski M. Stasiak, J. Wiewiora. Analytical modelling of the iub interface in the umts network. *Communication Systems, Networks and Digital Signal Processing, 2008. CNSDSP 2008. 6th International Symposium on*, pages 204–207, 2008.

- [83] R. Kreher. *UMTS performance measurement: a practical guide to KPIs for the UTRAN environment*. 2006.
- [84] F. Suili H. Shengye, Y. Wu and S. Hui. Coordination-based optimization of path bandwidth allocation for large-scale telecommunication networks. *Computer Communications*, 27:70–80, 2004.
- [85] y J. Wiewiora M. Stasiak, P. Zwierzykowski. Modeling and dimensioning of the iub interface in the umts network. *Ubiquitous Computing and Communication Journal. Special Issue CSNDSP 2008*, 2009.
- [86] E. Vivier M. Terre and B. Fino. Optimisation of downlink resource allocation algorithms for umts networks. *EURASIP J. Wirel. Commun. Network*, 2005:573–578, 2005.
- [87] Y. Ouyang y M.H. Fallah. Performance analysis for umts packet switched network based on multivariate kpis. *International Journal of Next-Generation Networks 2.1*, pages 79–92, 2010.
- [88] Hervé Dubreil Zwi Altman Rana Khanafer, Lars Moltsen and Raquel Barco. A bayesian approach for automated troubleshooting for umts networks. *The 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 06)*, pages 1–6, 2005.
- [89] D. Zhang S. J. Oh and K. M. Wasserman. Optimal resource allocation in multiservice cdma networks. *Wireless Communications, IEEE Transactions on*, 2:811–821, 2003.

Cálculo de bloqueo en redes multiservicio

- [90] J. Kaufman. Blocking in a shared resource environment. *Communications, IEEE Transactions on*, 29:1474–1481, 1981.
- [91] V. B. Iversen and S. N. Stepanov. The usage of convolution algorithm with truncation for estimation of individual blocking probabilities in circuit-switched telecommunication networks. *Proceedings of the 15th International Teletraffic Congress, ITC*, 15:1327–1336, 1997.
- [92] J.W. Roberts. A service system with heterogeneous user requirements application to multi-service telecommunications systems. g. pujolle, redaktor. *Proceedings of Performance of Data Communications Systems and their Applications*, pages 423–431, 1981.
- [93] F. Suili X. Xu, Y. Wu. Research on the calculation of call blocking probability in large-scale multimedia network. *Mobile Technology, Applications and Systems, 2005 2nd International Conference on*, pages 4–10, 2005.

Técnicas de aprendizaje e inteligencia

- [94] Rubén Ruiz García Concepción Maroto Álvarez, Javier Alcáraz Soria. *Investigación operativa: modelos y técnicas de optimización*. 2002.
- [95] Juan Ram Rabu Al Daniel Rivero, Marcos Gestal. *Genetic Programming. Key Concepts and Examples*. 2011.

- [96] Luis Joyanes Aguilar James A. Freeman, Rafael García-Bermejo Giner. *Redes neuronales algoritmos, aplicaciones y técnicas de programación*. 1993.
- [97] Toby Segaran. *Collective Intelligence*. 2007.

Conceptos matemáticos y estadística

- [98] Harald Cramér. *Random Variables and Probability Distributions*. 2004.
- [99] Charles Kingman John Frank. *Poisson Processes*. 1993.
- [100] Peter Olofsson. *Probability, Statistics, and Stochastic Processes*. 2011.

Regulación automática

- [101] Eduard Beltrán Alberti. *Procesado Digital de Señales. Fundamentos Para Comunicaciones y Control*. 2006.
- [102] Univ. Pontificia de Comillas F. Luis Pagola, Luis F. Pagola. *Regulación Automática*. 2006.

Herramientas informáticas

- [103] Gabriel Valiente Feruglio. *Composición de textos científicos con Latex*. 1997.
- [104] Dr. José Antonio Macías Iglesias. *Matlab. Una introducción con ejemplos prácticos*. Wiley, 2006.
- [105] Bernardo Cascales Salinas. *El libro de Latex*. 2003.
- [106] J. Gabriel García Tomás Aranda. *Notas sobre Matlab*. 2002.